

France's research base needs a shaking

After years of failed reforms, the time is ripe for the French government to loosen rigidities within academic institutions and the CNRS. Top of the list are the terms of employment of young researchers.

The collective sense of gloom and doom — “*la morosité*” — that has hung over France for the past decade seems to be lifting under the Prozac-like effect of the socialist government of Lionel Jospin. Scientists in particular are relieved to have escaped brutal cuts in public spending. Instead, the socialist victors have made education and research a priority, and increased the budget for civil research and development despite the difficult task of keeping the public deficit below the 3 per cent threshold needed to join the European single currency. A bigger psychological boost has come from the government's decision to create thousands of new full-time research posts in the research organizations and universities, ending years of stagnation in recruitment of young scientists.

The government's clear support for science has made scientists and their trade unions more open to the prospect of change. Claude Allègre, the education and research minister, has therefore bought himself a window of opportunity to bring about reform. Moreover, demographic trends indicate that the size of the student population is about to plummet, after decades during which universities were pre-occupied with catching up with exponentially growing student numbers. As Allègre aptly asserts: “France has won the battle of quantity, now it is time to win the battle of quality”.

Catherine Bréchnac, whom Allègre recently appointed as director general of the CNRS, last week announced an increase of almost one-tenth in funding for laboratory research, by taking money from big national physics facilities and strategic programmes (see page 895). This reflects Allègre's conviction that the balance between support for investigator-driven research and national wealth-creation strategies should favour the former, a policy that runs counter to the top-down tendencies elsewhere.

Although such moves will improve the lot of many scientists at the bench, they will not resolve the underlying structural problem repre-

sented by the abysmally small sum that the CNRS is able to devote to research grants, given that its salary bill eats up more than three-quarters of its total budget. Bréchnac's plans to increase competition for research grants — if they materialize — would mean that the sums available are spent more effectively, but it is clear that France could run a more productive and flexible research base by having fewer tenured scientists sharing the research-grant pie.

French scientists enjoy civil servant status: in other words, permanence that is enviable but which stifles competition, reduces mobility among researchers — there is no system of postdoctoral fellowships in France — and creates bloated and slow-moving bureaucracies. Neither Allègre nor many other French seem to want to tamper with the basic premise of life tenure. That principle (so long as it is sustainable) necessitates greater *dirigisme* on his part in enforcing mobility between different components of the government-funded research base. It is increasingly indefensible that researchers recruited to the CNRS should be guaranteed a lifetime career there. Better, surely, to change the terms of employment of new recruits so as to enable them to spend more of their later years teaching within the universities. Similarly, a robust system of postdoctoral fellowships could provide an efficient halfway house to permanent employment.

Grandiose policies to promote mobility have repeatedly failed, as can be judged by the fact that only a handful of researchers move each year between the research organizations, universities and industries. Will Allègre do any better? He knows well the full range of problems facing French research, and seems committed to implementing remedies through a policy of negotiated change rather than brusque upheaval imposed from above. Such an approach looks wise in exploiting researchers' current good will. But it should not become an excuse for inaction or half-measures in the face of the inevitable resistance to a much-needed shake-up. □

University reform overdue

Japan's higher education institutions cannot carry on business as usual.

Leaders of Japan's best universities have long maintained that their biggest problem is simply a lack of money compared with their counterparts in the West. If only the government would give them more funds, it is said, they will turn themselves into ‘centres of excellence’ in research and teaching through ‘self-evaluation’.

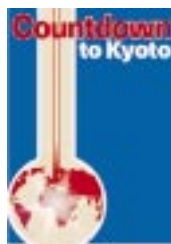
Nothing could be further from the truth. Although Japan's best universities do harbour some brilliant researchers and teachers, they, like French institutions (see above), suffer from intellectual sclerosis caused by an existence that is too comfortable by half. They also need greater freedom from the bureaucratic strait-jacket placed on them by the education ministry and university administrators — who are often seconded from the ministry to ensure that the university stays in line with government policy.

The administrative reform council of the ruling Liberal Democ-

atic Party has made a radical proposal to meet the need for reform by converting some of Japan's national universities into more independent and accountable ‘agencies’ (see page 897). But the proposal has been quickly stamped on by the universities, the education ministry that oversees them, and Akito Arima, one of Japan's principal proponents of reform and external review. The universities are clearly concerned that such a move would deprive them of guaranteed government funding and will open them to public scrutiny, revealing the enormous waste that goes on in the system. There is also a genuine concern that Japan's universities are simply not ready for such drastic reform — they might be unable to cope with their new-found freedom. But if Japanese universities are truly to become world-class centres of excellence, they must accept healthy public scrutiny while taking on more responsibility for their financial affairs. □



Clinton struggles to appease all sides over climate proposals



[WASHINGTON] US President Bill Clinton sought last week to bridge the yawning gap between two audiences for his climate change policy — the international community and the US Senate — by doing what he does best, and offering a vague compromise designed to appease both groups.

On the face of it he failed: Clinton's proposal to reduce greenhouse gas emissions to 1990 levels by 2008–12 managed to attract sharp criticism from senior European officials and prominent senators, as well as the main environmental and industrial groups. But politically he may yet succeed in the classic Clinton manner — by occupying the hallowed middle ground on the issue, and leaving the difficult choices to his successors.

There are two key elements to the sketchy domestic policy proposal delivered by the president on 22 October. One is an early action plan, which includes some incentives to cut greenhouse gases but which largely relies on voluntary actions to cut emissions.

The second is an emissions 'trading scheme', ill-defined at this stage but probably international in scope and based on a successful US experiment in trading sulphur dioxide emissions, to take effect in 2008.

The success of the former element will depend on faith in the latter. "It's difficult to see how the early action plan is different from the existing voluntary arrangements," says one congressional staff member. The difference ought to be that when the trading regime arrives in 2008, polluters will receive credits for cuts they make now. But, as the staff member puts it, "politics being politics, they might choose to do nothing and lobby like hell" against the trading scheme actually being implemented in 2008.

Clinton administration officials seem to concede that tough choices will be left to others. "Each Congress is going to look at what's working and what's not working," says Gene Sperling, Clinton's chief economic adviser.

The incentives that Clinton will provide for voluntary action will be included in the 1999 budget (for the year beginning 1 October 1998), which he will present in February, Sperling says. The value of tax incentives and research and development spending amounts to \$5 billion over five years; a large sum, but not much in its potential impact on America's \$5,000 billion-a-year economy.

The research and development package is likely to reflect a recent report from a panel of

the President's Council of Advisors in Science and Technology chaired by John Holdren of Harvard University. Holdren called for an extra \$1 billion a year for energy research — only enough, he admits, to restore spending to its level of five years ago. Half the money would go to research on energy efficiency and the rest would be split between fusion, fission, coal and renewables. But the Republican-controlled Congress is likely to resist extra spending in these areas.

The other specific domestic action promised by Clinton is the restructuring of the electricity industry to enhance competition and cut carbon dioxide emissions. That, too, may be a tall order in the Congress, where legislation to deregulate the industry is likely to be drafted in a way that protects operators of existing coal-fired power plants.

But the main focus of attention at the conference of the climate convention in Kyoto, Japan, in December will fall on what the United States says it can achieve, rather than what it can actually deliver. In this regard, environmentalists were angry at a last-minute decision by the president to drop a pledge to cut emissions by 5 per cent by 2017. Carl Pope, president of the environmental group the Sierra Club, said he was particularly disappointed by that decision.

This retreat may leave the United States with some room for manoeuvre in the Kyoto negotiations. The United States may concede a little ground to the rest of the world on emission targets, in exchange for something that will satisfy its requirement that, as Clinton put it last week, "key developing nations meaningfully participate in this effort".

The concessions that the United States seeks include agreement by these countries



Clinton: may be able to leave difficult choices on carbon emissions to his political successors.

to participate in "joint implementation", which would allow rich countries to meet their targets by helping with emissions cuts in poor countries. Of the important players, only Argentina has so far agreed to this. Also anticipated in December is a 'Kyoto mandate' attached to the treaty, to bring developing countries into the process. Colin Macilwain

Australia eases path to Kyoto agreement

[LONDON] The British government has brokered a deal between Australia and developing countries in which Australia has agreed to drop its opposition to a greenhouse gas reduction protocol in return for a commitment to "significant reductions" from developing countries at a later date.

The deal was struck after a series of meetings between foreign ministry officials at the Commonwealth heads of government conference in Edinburgh, which ended on Monday (27 October).

But Australia's position remains ambiguous. Shortly after agreeing to the declaration, Prime Minister John Howard

appeared on television claiming that the document amounted to a triumph for Australia, as it did not imply a commitment to legally binding emissions reductions.

Australia is opposed to these on the grounds that they would hurt the country's coal and electric power industries. It fears that jobs will be lost if Australia's power companies relocate to neighbouring South-East Asia to take advantage of more liberal rules on greenhouse gases.

Developing countries, on the other hand, do not want to reduce their emissions just yet. Some, such as India and China, are concerned about the possible adverse impact on their nascent industries. Others ▶

► believe that premature discussion of developing country obligations could detract from the main purpose of the Kyoto conference, at which developed countries are due to sign a legally binding emissions reduction protocol.

Climate change was not on the initial agenda of the Commonwealth conference. But British officials realized that the meeting would be an ideal occasion to try to generate consensus between Australia and other Commonwealth countries.

Another aim was to reduce the gap between developing countries and the United States, which confirmed last week that it will commit itself to a legally binding target to stabilize emissions at 1990 levels between 2008 and 2112, but only if there is 'meaningful participation' from key developing countries (see previous page). Argentina has in principle agreed. But other countries — including Brazil, China and India — so far remain opposed.

Environmental groups are also predictably angered, although their rhetoric has lacked some of its previous fire. "President Clinton's watching a house starting to burn. He wants to do something about it, but chooses to pull out the garden hose when he really needs a water truck," says Jennifer Morgan of the US Climate Action Network, based in Washington DC.

Some groups were optimistically expecting the United States to announce slightly more ambitious targets — perhaps a reduction in emissions to below 1990 levels — but coupled to much tougher conditions. Others, however, were privately relieved that Clinton confirmed his commitment to a legally binding target, rather than "just do nothing".

Last week, the Group of 77 (G77) nations and China endorsed the European Union's target to reduce greenhouse gas emissions by 15 per cent by 2010. One delegate from a major industrializing country says that this will remain the G77's "bottom line" position. Another developing country delegate says that the United States will have to make concessions if there is to be agreement at Kyoto.

The G77's endorsement of the European position was due to an agreement between the Alliance of Small Island States and oil-exporting countries about the latter's demands for compensation for revenues lost as a result of the reduced demand for oil. Both groups of countries have now agreed to

call for compensation for any country that suffers damage from the effects of either climate change or climate change policies.

Meanwhile, discussions this week in Bonn between the United States and developing countries — the last formal negotiations before Kyoto — have been making little headway. A two-day informal meeting before the start of the formal talks on 20 October resulted in deadlock.

Nonetheless, efforts to bring the US and developing country positions closer will continue during this week's visit to the United States by China's president Jiang Zemin. Three days of direct talks hosted by Japan are also scheduled in the second week of November.

Ehsan Masood

Climate panel to expand its membership

[BONN] The United Nations' advisory panel of climate scientists is to be expanded, and its procedures restructured, according to Bob Watson, the incoming chair of the Intergovernmental Panel on Climate Change (IPCC).

In future, the panel will include more scientists from developing countries and economies in transition, and business and development groups, Watson told the UN climate meeting in Bonn last week (see above). Reviews of scientific literature will also include material in languages other than English.

IPCC's Third Assessment Report on the world's climate, and a summary 'Synthesis' report for policymakers, will be ready by 2001, he said. The main report will focus heavily on regional aspects of climate change, while the synthesis report will consider key policy-relevant topics. Proposed answers to these will be circulated to governments for comment before being approved by the IPCC.

The changes have been partly designed to reduce conflict between scientists representing governments and those belonging to

independent research establishments. The previous second assessment report, particularly the summary document, was often surrounded by controversy (see *Nature* 378, 524; 1995).

Policymakers wanted answers to specific policy-related questions in language they could understand.

But the scientists drafting the reports were reluctant to stray beyond the confines of science, whose findings were often uncertain or subject to debate, a position which confused policymakers.

E. M.

Controversy flares over AIDS prevention trials in third world

[WASHINGTON] A furious exchange between researchers at Johns Hopkins University (JHU) and public health advocates, carried out in letters to the Secretary of Health and Human Services (HHS), has fanned the flames of a growing controversy around the ethics of US government-sponsored trials seeking to prevent perinatal AIDS transmission in developing countries.

At issue is whether it is any longer justified to use placebo in these trials when an expensive but effective therapy has been found and is routinely used in women in industrialized countries.

On 23 October, Public Citizen, a Washington-based advocacy group, wrote to Donna Shalala, the HHS secretary, pointing out that researchers at the JHU School of Public Health in Baltimore, Maryland, had "tentatively" dropped plans for a placebo arm in an NIH-funded trial set to begin in Ethiopia as soon as February.

The trial, involving some 900 women, had originally proposed comparing short, less expensive regimens of an anti-AIDS drug, zidovudine (AZT), in pregnant women, against women given a placebo. The JHU researchers, Neal Halsey and Andrea Ruff, say that they have made "contingency plans" to modify the trial design by dropping the placebo arm "if other studies document the effectiveness of a practical short course regimen [of AZT] in a developing country".

The doctors at Public Citizen, Peter Lurie and Sidney Wolfe, jumped on this as evidence that the JHU researchers are "acknowledging that it is possible to conduct a scientifically valid and useful study without the use of a placebo arm". They demanded in their letter that Shalala immediately order US government-sponsored researchers "to stop any arm of their studies in which women are denied access to antiretroviral drugs, and to provide at least short-term

AZT for all women now getting a placebo."

The next day, the JHU researchers shot back a letter of their own to Shalala. "Our changes are not being made to accommodate any outside groups," they said. Rather it was in response to information expected early in 1998 from other United Nations and US-supported trials.

They accused Public Citizen of spreading "mistruths and distortions" in press releases and said Wolfe and Lurie "have deliberately misrepresented our position and actions in an attempt to undermine the ongoing trials" of AZT in developing countries. Wolfe, in turn, accused Ruff and Halsey of "cheap justifications of what they are doing."

A spokesman for Shalala, said on Saturday that Shalala would have no immediate response to the letters. But, he added "we're keenly aware of the ethical implications of these experiments and we are always looking at new information."

Meredith Wadman

CNRS wants more competition for grants

[PARIS] The French Centre National de la Recherche Scientifique (CNRS), Europe's largest fundamental research agency, is substantially to increase funding for its 1,300 laboratories. The rise will be paid for by cuts in spending on large science facilities and strategic research programmes.

The measures were announced last week by Catherine Bréchnac, the recently appointed director general of the agency. Speaking at her first press conference since taking office, she also confirmed the creation this year of 425 posts at the 25,772-strong agency — a marked reversal of the past few years' stagnation in recruitment.

Bréchnac's proposed reforms are modest compared with recent unsuccessful plans to reform the agency, ranging from dismantling it completely to transferring large numbers of its laboratories to the universities (see *Nature* **371**, 639; 1994). The research minister, Claude Allègre, is keen for CNRS to slash its bureaucracy and shift the emphasis towards investigator-driven research.

Under Bréchnac's plans, CNRS will increase the budget for spending by laboratories on equipment and running costs by 7–8 per cent this year, and will introduce greater competition for these funds. The increase has been warmly welcomed by researchers, who have been accustomed to living on a shoestring. The agency spends more than three-quarters of its budget on salaries, and 10 per cent on 'big science' facilities, which leaves little for everyday laboratory running expenses.

This year's CNRS budget of FF14.7 billion (US\$2.5 billion) includes a 2.2 per cent increase for salaries, but otherwise is up just 1.1 per cent — less than inflation. To pay for the increase in research laboratory spending, the agency intends to cut its FF445-million budget for big science facilities by about 10 per cent, and to reduce its FF160 million spending on strategic programmes by a similar proportion.

As a result, the national Saturn accelerator at Saclay, near Paris, will be closed at the end of the year. Funding for Ganil, a new heavy-ion accelerator, and other large physics facilities will be reduced.

Bréchnac hopes that the consequences of these reductions may be partly offset by greater foreign participation in national facilities. At the same time, she indicated her support for Virgo, a planned Franco-Italian gravitational wave detector, and Soleil, a proposed FF1 billion synchrotron, which has been frozen for the time being by the science ministry.

Bréchnac also promises greater competition in the distribution of research funds. At present, funds are distributed on a *pro rata* basis to laboratories, depending on their size

and overall performance. In future, much greater emphasis will be placed on the "originality and creativity" of individual teams and projects.

"We will be much more selective," says Jacques Sevin, CNRS director for strategy and programmes. He says that whereas strategic research goals, such as biotechnology, will not change much, these will increasingly be supported by "unsolicited proposals".

The procedures for evaluating research groups will become more international. Foreign scientists will be asked both to rank French research groups and proposals and to sit on the review panels that evaluate laboratories every four years.

Three-quarters of the money available for laboratories will be distributed directly on this basis. In one innovation, however, 10 per cent will be used to create a special fund to finance promising new projects or researchers. Another 15 per cent will be reserved for interdisciplinary projects in, for example, biotechnology, new materials, environment and telecommunications.

Many scientists have welcomed the promise to encourage greater competition for funding, although there is some scepti-



Bréchnac: more funds for running costs and equipment, more competition for project grants.

cism about how this will work out in practice. "Let's wait and see," says Pierre Chambon, head of the French Institute of Genetics and Molecular Biology, near Strasbourg.

The National Union of Scientific Researchers is giving Bréchnac the benefit of the doubt. Henri-Edouard Audier, a member of the union's board, says that although the details of Bréchnac's plans are still too vague to predict the outcome, she is a "pragmatic" person with whom the unions feel "they can do business".

Declan Butler

Nuclear agency may absorb physics institute

[PARIS] France's Centre National de la Recherche Scientifique (CNRS) may be about to transfer the entire National Institute for Nuclear and Particle Physics (IN2P3) — which accounts for one-tenth of the CNRS's staff and budget — to the atomic energy commission, CEA.

The move is intended to increase CEA's fundamental research capacity, which is widely considered to lag behind the commission's engineering prowess (see *Nature* **374**, 104; 1995). Long-term fundamental research is needed to provide solutions for managing nuclear waste and developing safer and cleaner reactors.

The proposal falls within government plans to clarify the responsibilities of the various research organizations, which have become confused. The CEA, for example, has diversified away from its core mission of supporting nuclear research, and now has large

groups in climatology, Earth sciences and life sciences. A series of transfers of laboratories from one agency to another is likely.

IN2P3 has 18 laboratories and 2,600 staff. It already has close links with CEA.

Although particle physics might at first sight have little to do directly with nuclear research, one IN2P3 official points out that the institute's laboratories have traditionally carried out both particle physics and nuclear research. Observers say the proposal to transfer IN2P3 would regroup CEA's existing laboratories with those of CNRS, and allow a common scientific strategy for nuclear research. IN2P3 would also benefit from acquiring CEA's groups in astrophysics and other areas of basic physics.

Some observers say it is essential that IN2P3 remains semi-autonomous, as otherwise it risks being absorbed within the engineering-oriented strategy

of CEA. Others worry that moving IN2P3 may weaken its links with the universities.

The transfer of the institute would represent the end of the CNRS policy of covering all areas of basic research, and would bring a marked reduction in the size of the agency.

The CNRS's domination of French research is also being challenged by the universities. According to Maurice Gross, CNRS's director of university relations, universities are playing an increasing role in the administration of CNRS laboratories located on their campuses. Research minister Claude Allègre is keen to see the universities having greater control over research.

Jacques Sevin, CNRS director for strategy and programmes, agrees that "there is a risk that the role of CNRS will be reduced". But he says this need not necessarily be detrimental to research.

D. B.

UK tightens regime for animal research

[LONDON] The UK government has agreed to strengthen the law on the use of animals in research. The number of inspectors will be increased, and scientists applying for a licence for an animal experiment will have to explain why a non-animal alternative is unsuitable.

But, despite a pre-election pledge by the Labour government to restore a previous cut, there is little new money to fund research into non-animal alternatives. Nor is there a promise to ban the use of animals in cosmetics testing, another pledge.

The decisions were announced last week by Jack Straw, the Home Secretary, in response to the publication of an interim review of the 1986 Animals (Scientific Procedures) Act by a committee of government advisers.

Animal rights lobby groups, as well as those representing scientific research, have broadly welcomed the review. But both sides have criticized the decision not to restore funding for research into non-animal alternatives to its level of two years ago. The sum was reduced to £182,000 last year by the then Conservative government. Labour has agreed to increase the amount by £20,000, still less than the £250,000 spent in 1995.

Les Ward, director of Advocates for Animals and a member of the Home Office advisory committee that reviewed the legislation, describes the extra funding as a "pittance", and Robert Coombes, scientific director of the Fund for the Replacement of Animals in Medical Research, says it is "abysmal". "Each major project needs around £100,000–£150,000," says Coombes.

Mark Matfield, director of the Research Defence Society, agrees that the government should spend more on investigating alternatives. He says that although most of the funding for alternatives comes from industry, the level of government support indicates "how much importance it attaches to alternatives".

The review recommends increasing the "public understanding of the principles behind cost-benefit assessments, the method by which the suffering of animals is weighed against public benefit". The sentiments behind the recommendation have been welcomed, although observers remain unclear about what it means in practice, and some remain sceptical about what it will achieve.

Maggie Jennings, director of the animals in research division of the Royal Society for the Prevention of Cruelty to Animals, says

the general public does not approach this issue in terms of costs and benefits.

Laurence Smaje, director of the medicine, society and history division of the Wellcome Trust, says that surveys show "most people think that experiments are done on cats and dogs to test cosmetics".

He says an explanation of the methodology of cost-benefit assessments will help individuals already engaged in the debate about the use of animals in experiments, but will be meaningless to others.

Smaje says that public understanding of the debate about animals in research requires broader initiatives aimed at different audiences, particularly schools and universities, but also involving museums and science centres. He says that this last group has been reluctant to give space to the issue of animals in research, from fear of being targeted by animal rights extremists.

Another challenge, Smaje says, is to persuade teachers who oppose the use of animals in research to encourage a balanced debate in classroom discussions. He also believes the question of animals in research should be discussed during undergraduate science courses at universities. Ehsan Masood

Astronomers review options after fire ravages gamma array

[MUNICH] Astronomers are assessing the damage caused by a fire that raged last week through the site of the High-Energy Gamma Ray Array (HEGRA) at La Palma in the Canary Islands (pictured right), only months before it was due to start operating at full capacity.

The array, a collaborative project run by laboratories in Germany, Spain and Armenia, detects and analyses cosmic rays and gamma rays hitting the Earth, and searches for galactic and extragalactic gamma-ray sources.

The fire appears to have been connected with the burning of scrub during landscaping activities in the national park in which HEGRA is situated. National park regulations stipulate that the site may not be cleared of scrub, and the tinder-dry gorse bushes burning between the detectors caused most of the damage.

HEGRA comprises a 200-square-metre chequerboard system of gamma-ray detectors, known as Cherenkov counters. These are punctuated by six imaging Cherenkov telescopes which are the most sensitive ground-based gamma-ray detectors in the world.

One-third of the array detectors and one of the telescopes were destroyed by the fire, at a cost estimated at between DM1.3 million and DM1.6 million (US\$730,000–\$900,000). Further damage to the



instruments' sensitive electronics by hydrochloride vapours released from burning plastic may be revealed later, according to Heinrich Völk, a director at the Max Planck Institute for Nuclear Research in Heidelberg, and a spokesman for the project.

Völk says the observatory was saved from complete destruction by scientists based at the site, who tried to bring the fire under control with extinguishers before the fire brigade arrived, as well as by a fortuitous change in wind direction.

In its various stages of completion, HEGRA had been operating for about six years. Four telescopes had been operating since the end of last year. Completion of the entire stereoscopic system of six telescopes was expected this winter.

HEGRA scientists are still able to work with the undamaged detectors and telescopes, but at greatly reduced sensitivity and efficiency. They are discussing possible compensation with insurance agencies. But because of laws restricting insurance cover of equipment bought with public money, the instruments' total value is not covered.

Also, as Völk points out, money is not the only problem when replacing damaged instruments. "They are not off the shelf," he points out. Laboratory staff who built them are not necessarily available to rebuild them.

The HEGRA teams will be asking for help from the agencies that have paid most of the costs of the multimillion dollar project — the German federal research ministry and the Max Planck Society.

Alison Abbott

HEGRA

Japan's universities resist 'agency' plan

[TOKYO] Japanese government proposals to transform the country's two leading universities — Tokyo University and Kyoto University — into independent 'agencies' have met strong resistance from both universities and from the Ministry of Education, Science, Sports and Culture (Monbusho).

A subcommittee of the Administrative Reform Council of the ruling Liberal Democratic Party (LDP) has proposed turning two national universities into independent administrative agencies (*dokuritsu gyousei hojin*). The proposal forms part of a drive for administrative reform launched last year by the prime minister, Ryutaro Hashimoto.

The universities at present come under the control of Monbusho. Under the plan, which was presented to a committee meeting last week, they would be given administrative independence, and a system would be developed for targeting funding at institutions and research groups with good track records. Initially it had been intended to focus on all 98 national universities, but this idea was discarded as impractical.

Monbusho's management of national universities has frequently come under fire for being inefficient and too rigid, and for dispersing funds to universities without adequate consideration of their performance. The government has also been criticized for a system under which research funds provided on an annual basis have to be used by the end of the fiscal year. As unspent funds cannot be carried over to the next fiscal year, many in universities feel that the system leads to taxpayers' money being wasted (see box).

Kiyoshi Mizuno, an LDP member and one of the key members of the Administrative Reform Council, argues that the changes would allow universities to be run more flexibly, because the number of civil servants who

oversee university administration offices would be reduced, and a wider range of funds would become available. Mizuno says that the plan will therefore help to raise universities' standards and make them "more competitive".

But Monbusho and the two universities claim that the proposal would jeopardize the quality of research and academic standards and lead to a significant increase in tuition fees. Hiroo Imura, president of Kyoto University, is worried that the agencies, which would be reviewed every three to five years, will "inevitably affect consistency in university education and research. This will lower our international competitiveness."

"Japanese universities are under great pressure to change, but we do not want the change to be imposed on us," Shigehiko Hasumi, president of Tokyo University, said at a press conference last week at which he expressed the university's opposition to the agency plan.

Hasumi says the plan should have been discussed at a national level, involving the universities and the general public, before being taken up in the political arena. "We are not entirely satisfied with how things are at present, but we would not make any reform unless we feel happy about it."

Akito Arima, president of the Institute of Physical and Chemical Research (RIKEN), and former president of Tokyo University, points out that Japan spends only 0.9 per cent of gross domestic product on higher education, whereas the United States and some European countries spend twice that.

"Japanese universities should be funded by the government — or receive more support at the least," says Arima, a member of the reform council. Arima represents the scientific community in the panel and is reported to have protested strongly against

the plan. He claims that as administrative agencies involved in profit-making activities, the objectives of universities would be focused on financial matters.

"This is the last thing universities should be doing," he says. "Instead, there should be a reform on more basic issues such as endorsing a proper peer review system to improve the quality of each institution."

The main concern of the universities, however, appears to be a fear that the government will cut their funding if they become agencies, forcing them to rely more on raising funds from other sources. Monbusho says that this could well happen in Japan's difficult economic climate, and that the universities might be left without a secure source of funding.

Mizuno, who has been advocating the agency plan, feels that both universities will eventually accept the main thrust of the plan, providing the details are thrashed out more fully. But he accepts that convincing Monbusho will not be easy.

Asako Saegusa

nature

is organizing a half-day conference on:

UK university research: Dearing and after

Friday, 7 November 1997

2.00 p.m. – 6.45 p.m.

Royal Society, 6 Carlton House Terrace, London SW1Y 5AG ENTRY FREE

The meeting will end with a panel discussion and be followed by a reception

Admission by ticket only, available from: Mary Sheehan, 4 Crinan Street, London N1 9XW or m.sheehan@nature.com Fax: 0171 843 4595/6/7

Speakers:

- **Sir Ron Oxburgh** (Rector of Imperial College, London, and a member of the Dearing Committee)
- **Mark Ferguson** (School of Biological Sciences, University of Manchester)
- **Jean-Patrick Connerade** (Department of Physics, Imperial College, London; member of the executive council, Save British Science)
- **Dame Bridget Ogilvie** (Director, The Wellcome Trust)
- **Brian Fender** (Chief Executive, Higher Education Funding Council for England)
- **David Triesman** (General Secretary, Association of University Teachers)

Inquiry investigates 'hiding' of unspent funds

[TOKYO] Five research-related faculties and four research institutes at Kyoto University are under investigation for allegedly 'hiding' unspent research funds. Such funds should have been returned to the government at the end of the fiscal year.

The Ministry of Education, Science, Sports and Culture (Monbusho) disburses funds to national universities in July. But money not spent by the beginning of the next fiscal year, on 1 April, must be returned by law.

Last week, university

administrators at Kyoto announced that tens of millions of yen have been hidden over a five-year period. Members of the departments under investigation — which include the faculties of pharmacology, engineering, science, and the institutes of chemistry and virology — are believed to have deposited unused money with several research equipment suppliers. These suppliers are alleged to have issued fake statements and kept the money until the next fiscal year.

Many researchers are unhappy about the inflexibility of the current arrangement, which means that departments often have either too little or too much money. "If we do not use up the allocated money, our budget might get cut," says a professor from one of the institutes under investigation.

Administrators say that the 'hidden' money, first discovered in April this year, "will be returned to the government, depending on the result of the investigation".

A. S.

Genetic testing for Alzheimer's disease 'not appropriate'

[PALO ALTO, CALIFORNIA] Despite advances in the understanding of the genetic basis of Alzheimer's disease, genetic testing for the disease is not appropriate for most people — even for diagnostic purposes. That is the conclusion of a group of scholars, physicians, lawyers and patient advocates brought together by Stanford University's Program in Genomics, Ethics and Society.

In a draft report released last Saturday (25 October), the group points out that the value of the additional information that genetic testing provides must be balanced against the financial and psychological cost. Predictive tests should be considered in the context not only of their sensitivity and specificity, but also in relation to prospects for treatment and the non-medical risks and benefits of the test.

Only those with a clear family history of Alzheimer's disease, especially with onset before the age of 60, might be possible candidates for testing for highly penetrant mutations, such as *APP*, *PS1* or *PS2*, the group says. Most individuals are unlikely to have these genes, which are estimated to cause less than 2 per cent of Alzheimer's cases.

More common is a susceptibility gene, *APOE*, of which one allele, *E4*, appears to be linked to about a 14-fold increase in risk for late-onset Alzheimer's among certain ethnic groups. About 50 per cent of all Alzheimer's cases may be attributable to *E4*, says Neil Risch, a professor of genetics at Stanford and co-author of a recent metanalysis. As a result, a test for *APOE* has much wider applicability and greater commercial potential than the other mutations.

But there is general agreement that *APOE* offers little predictive value — partly because there are so many causes of Alzheimer's, and not much is known about how the gene acts. Nor, according to the Stanford group, is a diagnosis confirmed by the *APOE* allele of clear value, especially in the light of its lack of usefulness for treatment, and its wide-ranging implications for family members who must act as surrogates for the person with dementia.

The group also pointed to the significant financial cost, which can be as high as \$1,500 to \$2,000, and to the potential psychological and social harm to early-stage affected individuals and relatives, who would inevitably learn of the probable nature of their own genotype.

A diagnostic test for *APOE* was first offered in 1993, but was withdrawn three months later after criticism. Now Athena Diagnostics offers tests for both *APOE* and *PS1*, but only for diagnostic purposes after a doctor has signed a statement that the gene donor is clinically demented. Sally Lehrman

NIH pilots faster feedback for grant resubmissions

[WASHINGTON] J. Brice Weinberg, a professor of medicine at the Veterans Affairs & Duke University Medical Centers in Durham, North Carolina, had that sinking feeling early last March when he received an e-mail informing him that he had failed to obtain funding for a major malaria study.

But his disappointment turned to hope when he read a paragraph at the end of the rejection from the US National Institutes of Health (NIH). This said his application could be quickly reconsidered if he answered a list of questions in an 'abbreviated amended' application.

By the end of April, NIH's Tropical Medicine and Parasitology (TMP) study section had Weinberg's seven-page response. And by September he had received the first instalment of a \$2-million, five-year grant for an international, multi-centre study of the function of nitric oxide in severe malaria.

The speedy reapplication gave him and his co-researchers "a terrific boost", says Weinberg. Not only were they saved the effort of rewriting the application for standard resubmission, but they saw their money at least four months earlier than would have been the case if they had been successful on a conventional second try.

The speedy reprocessing is just one part of an experiment that could eventually be introduced widely across NIH in which peer reviewers in the TMP study section are trying to reduce the time from the receipt of applications to when scientists see money — or learn why they are not getting any.

It's a "total package" that is trying to use new technology to improve the review process, says Elvera Ehrenfeld, who heads NIH's Center for Scientific Review — formerly the Division of Research Grants. For example, the trial programme also aims to return assessments electronically to near-miss applicants within days of a study section meeting, rather than the typical six to eight weeks.

"Not only do those who get funded see a benefit, but those who don't get funded but are on the borderline, also benefit," says John McGowan, the pilot programme's overseer. McGowan is deputy director of the National Institute of Allergy and Infectious Diseases (NIAID), which funds almost all TMP grants.

Moriya Tsuji, an assistant professor of parasitology at New York University School of Medicine, says that receiving reviewers' criticisms in mid-July 1996 rather than mid-August enabled him to alter preliminary experiments and make critical revisions to his application in time for a 1 November resubmission deadline.

"If I had started in the middle of August, I

wouldn't have finished [by the deadline]," says Tsuji, who on his resubmission received a \$650,000 grant to study the use of adenovirus as a vector for malaria vaccine.

The pilot programme also uses electronics to help reviewers. Study section members log on to a secured Internet website one week before they meet, and upload their assessments of applications. The chance to digest the written opinions of their co-reviewers before the meeting means that applicants get more considered, better reviews, says Diane Wallace Taylor, a professor of biology at Georgetown University, who until recently was a TMP reviewer.

But most important in speeding the process is that the study sections' critiques and scores are sent electronically within days directly to the NIAID Advisory Council, which must approve all grants. Top-scoring grants can therefore be immediately cleared by council delegates, shaving two to four months off the approval process.

The pilot programme — which ends next February — has raised some concerns. Franklyn Prendergast, director of the Mayo Cancer Center in Rochester, Minnesota, worries that the speedy amendment process could swamp both reviewers' time and institute budgets. And Tony Mazzaschi, a research policy analyst at the Association of American Medical Colleges, says: "The real challenge is going to be to do it in a much larger set of study sections."

But parts of the programme are spreading. The NIAID has since 1995 been using electronic approval by council for almost all institute-funded grants that fall within funding range. And the National Institute of Dental Research, the National Institute of General Medical Sciences and the National Institute of Deafness and Other Communication Disorders have all launched or are launching systems using electronic initial peer review.

Meredith Wadman



Does low cost mean low-value missions?

Tony Reichhardt

Despite an increase in the number of planetary missions, views are mixed about whether NASA's policy of 'better, faster, cheaper' is working.

Scientists have not all been convinced about the effectiveness of the 'leaner' approach to planetary exploration — in the wake of the Cassini launch to Saturn, the selection of two Discovery missions and the Pathfinder landing on Mars.

Five years ago, the US National Aeronautics and Space Administration (NASA) introduced its Discovery line to bring down the price and complexity of planetary missions. Discovery projects are capped at about \$280 million and a maximum development time of three years.

Six Discovery spacecraft have now been approved for launch, after last week's selection of the Comet Nucleus Tour (CONTOUR) to visit near-Earth comets and the Genesis mission to return samples of solar wind particles. NASA hopes to choose one or more new missions every 18 months.

Conventional wisdom is that large, expensive projects such as Cassini (total US investment of \$2.6 billion) have become politically and economically impossible. That leaves Discovery and a separate line of Mars missions as the only approved US planetary projects for the future.

Yet these alone do not constitute a full, balanced programme of Solar System exploration, some scientists argue. Discovery has unquestionably increased the number of planetary spacecraft in the launch queue. But the scientific return on these missions has not been demonstrated, and some of their drawbacks are already apparent.

What priority for science?

So far, the only completed mission has been the \$200-million Mars Pathfinder, which was primarily a demonstration of new technology, with science as a secondary goal. It was an engineering and public relations triumph, and an important demonstration of lean mission management — all of which scientists applaud. At the same time, says one, it contributed "damn little" new scientific information.

The past existence of water on Mars was already well known. Pathfinder's most intriguing results remain ambiguous, partly because of the limited capability of low-cost instruments. The onboard rover, for example, photographed what appear to be conglomerate rocks formed by water, but the image resolution is too low to settle the

matter conclusively. The same goes for hints of quartz and other Earthlike minerals in Martian rocks: chemical data without mineralogy are unlikely to nail down the answer. In short, says one scientist who still praises Pathfinder as an engineering and management experiment, "our perceptions of Mars haven't changed that much".

Discovery missions also appear limited in the kinds of objects they can study. Officially, the programme is open to any destination in the Solar System. But of the six missions selected so far, four will target nearby comets, asteroids or the solar wind, which are relatively easy and inexpensive to reach.

The \$63-million Lunar Prospector, the cheapest concept yet selected, has experienced another kind of problem, also related to low cost — repeated delays owing to its reliance on an unproved launcher, which has suffered teething troubles.

Scientists interested in the small bodies of the Solar System may now be getting their just rewards, after enduring NASA's cancellation of several planned asteroid and comet fly-bys over the past two decades. But it is not yet clear that Discovery missions can afford to go anywhere else.

The three losers in the most recent round of competition were a spacecraft to orbit Mercury, another to return samples of the Martian moons Phobos and Deimos, and a third to study the atmosphere of Venus. The first two were priced at the high end of Discovery's \$280-million cost cap, whereas CONTOUR came in at \$154 million and Genesis at \$216 million. It is widely believed in the planetary science community that the cheapest Discovery proposals tend to win.

Christopher McKay of NASA's Ames Research Center — who admits that his own unsuccessful Discovery proposal to send a spacecraft to Jupiter's moon Europa might have broken the cost cap — says that the cost constraints mean it is appropriate to send these spacecraft to comets and asteroids.

Because instrumentation is limited on these cheap flights, it makes sense to pick destinations that have not yet been explored, where one or two focused experiments can advance scientific knowledge, says McKay. "Perhaps the only good science you can do with these missions is to go to new places."

But that means some of the Solar System's most intriguing destinations, including the



Pictures versus data? Pathfinder produced great images of Mars, but little scientific information.

innermost and outermost planets, may be out of Discovery's reach. The more powerful launch vehicles, nuclear-powered batteries, thermal controls and other extras that they require eat up most of a mission's cost, with little or nothing left for experiments.

NASA, however, is aware of the problem. The agency's plan — which most planetary scientists support — is to develop technologies it hopes will cut mission costs. Spending has been increased on developing lightweight spacecraft parts, efficient power sources, solar electric propulsion and other enabling technologies, particularly those required for deep space exploration.

Aiming for the outer planets

NASA also hopes to begin a new line of outer planets missions in fiscal year 2000 to fill the void left by Discovery. They would presumably cost more, but how much more is not clear. The top contender is a Europa Orbiter, with a solar probe and a mission to Pluto and the Kuiper belt also under consideration.

But the Pluto project has been struggling for several years to meet strict cost targets set by the NASA administrator, Daniel Goldin. Engineers at the Jet Propulsion Laboratory have so far been unable to meet those targets.

Jonathan Lunine of the University of Arizona, a Cassini investigator who also led a science definition team for the Pluto Express concept, says that doing meaningful science at Pluto for \$250–\$300 million continues to be "very challenging".

Most US planetary scientists accept the need to bring down the cost of planetary exploration, and applaud NASA's drive for economy. But McKay says: "People may have prematurely said there will be no more big missions." It may be, he adds, that "dollar for dollar, Cassini is doing a lot more science" than the Discovery missions will.

Lunine says: "If we're going to explore the Solar System in a scientifically and culturally exciting way, we're going to have to pay for it." In other words, there should be a mix of small and intermediate missions. And perhaps 'expensive' should not always be considered bad. □

Visiting researchers in France promised 'fast-track permits'

[PARIS] Foreign researchers invited to work in France are to be spared the long queues and endless paperwork associated with extracting a *carte de séjour* (residency permit) from the police, under plans announced this week by Claude Allègre, the research minister.

Allègre said that many leading researchers had complained to him about what many see as the humiliating and ludicrous delays and obstacles that often need to be overcome to comply with France's bureaucratic and strict immigration laws. Several had indicated that as a result they would not return to work in France again.

Allègre promised that all visiting researchers will in future benefit from a fast-track procedure, and will be supplied with the necessary documentation before leaving their home countries. Foreign students will also benefit from simplified procedures.

Ford links engineering and environment at MIT

[BOSTON] **Ford Motor Company announced last week that it is to give the Massachusetts Institute of Technology (MIT) at least**

\$20 million over the next five years to support the study of engineering design and on-the-job education for engineers. The money will also establish an MIT-directed consortium — drawing on researchers from the academic community, industry and government agencies — to address environmental issues.

According to MIT's president, Charles Vest, the "global challenges" of the future "require closer, more cooperative interactions among universities, corporations, and governments". Ford's vice-president, John McTague, says that "this new joint effort will help set a model for how universities and industries can work together to achieve mutual benefits".

Congress critical over Gulf War illness

[WASHINGTON] The Pentagon and the Department of Veterans Affairs should lose their power to investigate Gulf War illnesses, according to a report by a US House of Representatives committee that is expected to be publicly adopted today (30 October). The report by the House Government Reform and Oversight Committee, an advance copy of which was obtained by the *New York Times*, says that the two departments have handled investigations of Gulf War illnesses so poorly that

responsibility for further investigation should be moved to a separate federal agency — a suggestion opposed last month by William Cohen, the Secretary of Defense.

The report, which has been 20 months in the making, asserts that "toxic agents", including Iraqi chemical weapons, were likely to be responsible for health problems in thousands of veterans. Congressman Christopher Shays (Republican, Connecticut), who chaired the subcommittee that drafted the report, writes that investigations by the two departments have been "irreparably flawed" and "plagued by arrogant incuriosity". The Pentagon declined to comment before receiving the report.

Russia renews promise to pay scientists

[MOSCOW] **The Russian prime minister, Viktor Chernomyrdin, last week promised that the government would rapidly pay the total of 2.4 billion rubles (US\$400 million) owed to scientists who work for government research institutions, and will pay all of the money owing on the science budget by the end of October.**

Opening a meeting in the House of Government in Moscow on 24 October, Chernomyrdin said that the 1998 science budget will be 13.5 billion rubles, 3.85 per

cent of all state expenditures and close to the promised 4 per cent. Of this total, 4.8 billion rubles will be for fundamental research, 5.6 billion for innovative technologies, and 3 billion for space exploration.

Vladimir Bulgak, the vice-prime minister responsible for science, told the meeting that the government had changed its policy about financing science. He said it would no longer give money "to names, titles and signboards", but to the projects and programmes selected through competition. "There are now 4,364 Russian 'scientific' organizations, some of which have nothing to do with science, and many of which duplicate the work of others. Naturally, we are not going to support all of them."

Ex-chief of Los Alamos to head FBI laboratory



[WASHINGTON] Donald Kerr (left), the former director of the Los Alamos nuclear weapons laboratory in New Mexico, has been named as the new director of the Federal Bureau of

Investigation (FBI) Laboratory, which examines hundred of thousands of pieces of evidence each year for police forces across the United States.

The 700-strong laboratory has been under pressure to improve its performance after allegations that it had mishandled samples contaminated by explosives. Kerr, a physicist, says that his top priority would be to obtain accreditation from the American Society of Crime Laboratory Directors, and then the transfer of most of the staff from Washington to a new laboratory under construction at Quantico, Virginia.

'Alternative' Nobels for anti-nuclear activists

[LONDON] Two anti-nuclear activists, Mycle Schneider of France and Jinzaburo Takagi of Japan, are among the five recipients of this year's Right Livelihood Award, sometimes known as the 'alternative' Nobel prize. Schneider, a Paris-based activist who has written widely about the shipment of used nuclear fuel, and Takagi, an associate professor of nuclear chemistry at Tokyo Metropolitan University who has been sharply critical of Japan's plutonium policies, have worked jointly in opposition to the fast-breeder programmes in their respective countries.

Schneider and Takagi share the award, worth a total of US\$240,000, with Cindy Duehring, a US anti-toxic chemicals activist, Michael Succow of Germany, who has campaigned for the creation of nature

reserves in both the former East Germany and Soviet Union, and Joseph Ki-Zerbo from Burkina Faso, an historian of Black Africa.

New journal launch: *Nature Neuroscience*

The publisher of *Nature* is to launch next May a new monthly journal, *Nature Neuroscience*. Calls for papers will be issued within the next few months.

Over the past five years, three monthly journals have been launched — *Nature Genetics*, *Nature Structural Biology* and *Nature Medicine* — joined in 1996 by the relaunched *Nature Biotechnology*. All operate with editorial independence from each other and from *Nature* (see author information in www.nature.com for more details of the relationships). All have established themselves as the leading journals within their disciplines, without compromising *Nature's* continuing role in publishing papers of unusually broad interest from any discipline.

Surveys of neuroscientists have revealed a strong enthusiasm for a *Nature* journal publishing across all areas of fundamental neuroscience. Editorial staff vacancies are advertised on Classified pages 8 and 21 of this issue.

Philip Campbell
Editor, *Nature*

Recent trends in the BSE epidemic

Sir — The confirmation of 21 cases of new variant Creutzfeldt–Jakob disease (vCJD) in humans by 31 August 1997 and the continuing ban on the export of beef and cattle from the United Kingdom have stimulated a continuing interest in the past, present and future course of the bovine spongiform encephalopathy (BSE) epidemic in the United Kingdom. We report analyses of the latest available BSE case data (for which we are grateful to J. Wilesmith and colleagues at the Central Veterinary Laboratory) and discuss their implications for future trends in the light of recent culling programmes.

By 21 August 1997, 168,578 cases of BSE had been confirmed in Great Britain (GB). Analyses of the maternal cohort study^{1–5} and the GB case database⁶ have revealed an enhanced risk of BSE in calves born to infected dams, with the cohort study suggesting a maternal transmission rate of approximately 10 per cent over the last 5 months of the maternal incubation period. Analysis of dam–calf pairs in the case database revealed an enhanced risk in calves born up to 2 years before the onset of clinical signs of BSE in the dam with the greatest risk in those born after onset⁶.

The observed number of cases onset in 1996 (7,417) is in close agreement with the numbers predicted in recent analyses of the epidemic (7,386 with a 95 per cent prediction interval (PI), 6,541–8,856) (Figure 1) (ref. 7). So far, 1,659 BSE cases have been confirmed with clinical onset in 1997. Extrapolating from the cases already seen in 1997 and adjusting for the delay between reported clinical onset to confirmation by pathology, 3,592 cases are estimated for 1997 — consistent with model predictions (95 per cent PI, 3,006–7,664) (ref. 7).

Two control programmes based on culling of GB cattle have been implemented over the past 18 months, namely, culling of animals over 30 months of age (OTMS) introduced in May 1996 and a selective cull

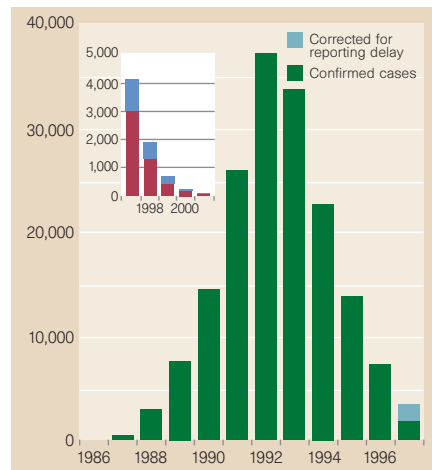


Figure 1 Number of BSE cases by year of clinical onset. The number of cases onset in 1997 was adjusted for confirmation delay by month for January to June and the monthly incidence for July to December was set equal to the adjusted number for June. An adjustment by month to the number of cases onset in 1996 increased the number from 7,417 to 7,441. Inset: Predictions of cases by year for 1997 to 2001 with no selective cull (blue) and assuming the Ministry of Agriculture selective cull was fully implemented by 1 January 1996 (red) (ref. 7).

initiated in March 1997, targeting herds with a history of BSE in line with the Florence agreement made with the European Commission in June 1996. Under the OTMS, 1,403,619 cattle had been slaughtered by 17 August 1997, but this number is only approximately 10 per cent greater than the expected number of deaths that would have taken place before the introduction of OTMS (based on the GB cattle survival distribution estimated using pre-1996 data^{7,8}). By 22 August 1997, 19,271 cattle had been slaughtered under the selective cull agreed with the European Commission. Estimating the impact of both schemes is made difficult by the absence of data on the ages of animals slaughtered under OTMS and limited

information on the animals targeted in the selective cull. However, Figure 1 provides approximate upper and lower bounds for future years.

A recent suspected case of vCJD in an individual who appears not to have consumed meat and meat products since 1985–86 has triggered further interest in the early stages of the BSE epidemic. Although the first case of BSE in cattle was diagnosed in 1986 (ref. 9), we estimate that up to 54,000 infected animals were slaughtered for human consumption before clinical onset of BSE between 1980 and 1985, depending on the level of under-reporting assumed¹⁰. Overall, before the introduction of the specified bovine offal ban in 1989, 480,000 infected animals were slaughtered for consumption¹⁰. Most of these animals were in the early stages of the incubation period. Since the incubation period distribution of vCJD is unknown, it is not possible at present to produce meaningful predictions of future trends in vCJD cases. This uncertainty will continue for the foreseeable future.

C. A. Donnelly

A. C. Ghani

N. M. Ferguson

R. M. Anderson

Wellcome Trust Centre for the
Epidemiology of Infectious Disease,
Department of Zoology,
University of Oxford,
South Parks Road,
Oxford OX1 3PS, UK
e-mail: christl.donnelly@zoo.ox.ac.uk

1. Wilesmith, J. W. *et al.* *Vet. Rec.* **141**, 239–243 (1997).
2. Donnelly, C. A. *et al.* *Appl. Stat.* **46**, 299–304 (1997).
3. Donnelly, C. A. *et al.* *Appl. Stat.* **46**, 321–344 (1997).
4. Curnow, R. N., Hodge, A. & Wilesmith J. W. *Appl. Stat.* **46**, 345–349 (1997).
5. Gore, S. M. *et al.* *Appl. Stat.* **46**, 305–320 (1997).
6. Donnelly, C. A. *et al.* *Proc. R. Soc. Lond. B* (in the press, 1997).
7. Anderson, R. M. *et al.* *Nature* **382**, 779–788 (1996).
8. Donnelly, C. A. *et al.* *Phil. Trans. R. Soc. Lond. B* **352**, 781–801 (1997).
9. Wells, G. A. H., *et al.* *Vet. Rec.* **121**, 419–420 (1987).
10. Ferguson, N. M. *et al.* *Phil. Trans. R. Soc. Lond. B* **352**, 803–838 (1997).

Rules needed on authorship

Sir — You report (*Nature* **389**, 105; 1997) that Dr Friedhelm Herrmann's lawyers argue that as senior author he would not have had any responsibility for possibly fabricated data and, moreover, that a full professor would not have a motive to commit fraud. Both statements are ridiculous.

Any scientist has to publish, or reputation and grant-money will be seriously endangered. The only way to publish for scientists no longer doing bench-work

themselves is to guide research and be senior authors. A senior author must be responsible for the scientific soundness and honesty of a paper, or he or she should not be in this distinguished place. They get reputation from it after all, and the quality of a paper is often inherently inferred from the name and status of the senior author.

The minimum requirement for senior authorship should be that the paper has been read, understood and worked on by intelligently and critically discussing it with the people who did the actual bench-work. In a perfect world, the senior author will also have provided ideas and mental input

through the course of the experiments and have seen the problems and difficulties of the study. There are heads of laboratories who provide just that, and are therefore rightly named the senior authors. I do not think that — as is sometimes the case — being head of a department or institute, providing laboratory space, or simply giving permission to use costly equipment as such justifies senior authorship. Unfortunately, there are no real rules, but it is time that there were.

C. Esser

Medical Institute of Environmental Hygiene,
Auf dem Hennekamp 50,
40225 Düsseldorf, Germany

No red alert over conservation *Red Lists*

Sir — In a recent Commentary article (*Nature* **389**, 436; 1997), Nicholas Mrosovsky highlighted the continuing work within the Species Survival Commission's volunteer network to improve the transparency of the IUCN (International Union for the Conservation of Nature) *Red List* by compiling documentation detailing the scientific justifications for the 5,000-plus listings in the latest volume.

As a member of the Species Survival Commission (SSC), Mrosovsky is well aware that it is our intention to make the documentation substantiating the inclusion of a species in the *Red List* publicly available as soon as possible. It is our hope that this enormous task will soon be completed and posted on the evolving Red List Internet website. However, as the SSC is a network of volunteers who donate their time and effort, sometimes such initiatives take longer to accomplish than its more enthusiastic members would prefer.

We recognize that there is a great deal of interest in the documentation for listings such as *Eretmochelys imbricata* (the hawksbill sea turtle). We do not believe that *E. imbricata* belongs in the "data deficient" category, and plan to make available very soon the documentation supporting the current listing.

The IUCN *Red List of Threatened Animals*, soon to be joined by the *Red List of Threatened Plants*, has been a respected standard reference for many years, and we plan to keep it so for many years to come.

David Brackett

(Chair, IUCN Species Survival Commission)
c/o Canadian Wildlife Service,
351 St Joseph Blvd,
Hull, Quebec,
Canada K1A 0H3
e-mail: david.brackett@ec.gc.ca

Sir — Mrosovsky has expressed concern about the credibility of IUCN. But is the IUCN as a whole really to blame?

The IUCN assessment of the Cuban hawksbill turtle proposal at the tenth Conference of Parties to the Convention on International Trade in Endangered Species (CITES) was based largely on information provided by selected members of the voluntary IUCN-SSC Marine Turtle Specialist Group (MTSG). (From 1982 to 1990 I was secretary general of CITES.) When MTSG members at that conference were given the opportunity to discuss their concerns publicly with the Cuban delegation, the group demonstrated a

superficial understanding of the proposal, the scientific analyses upon which it was based and even the IUCN principles about sustainable use with which the proposal complies. Group members showed little interest in resolving problems with Cuba, and were clearly committed to advocacy against Cuba.

Why the MTSG should have adopted such a strong advocacy position against Cuba is unknown. The information provided for the IUCN assessment came mainly from US members of the MTSG, and the United States has a vested interest in Cuba being isolated economically. Perhaps the controversial issue of turtle excluder devices in the US shrimp industry makes it politically difficult to admit that some sea turtle species are not 'endangered'. The MTSG has been slow to adopt the concepts and philosophies of sustainable use, despite their increasing acceptance by other SSC specialist groups and by the IUCN itself.

Nevertheless, as Mrosovsky states, IUCN credibility with CITES has been adversely affected by the antics of the MTSG. More effective ways of filtering philosophical and political biases will need to be found if the IUCN assessments are to be treated seriously and with respect.

Eugene Lapointe

(President)
IWMC World Conservation Trust,
1127 King Arthur — Unit 303,
Dunedin, Florida 34698, USA
e-mail: iwmc@ibm.net

Methyl bromide not so bad

Sir — The News story "Ozone treaty 'must tackle CFC smuggling'" is seriously misleading when it states that "Molecule for molecule, methyl bromide is considered at least 50 times more destructive to the ozone layer than chlorine from CFCs"¹.

Because of oceanic uptake and removal by oxygen-hydrogen radicals in the lower atmosphere, only about 4 per cent of methyl bromide molecules released at the surface survive the upward journey into the stratospheric ozone layer. Second, the dominant CFCs (chlorofluorocarbons) have two or three chlorine atoms per molecule. Further, only about 35 per cent of atmospheric methyl bromide can be shown to be under human control². Therefore, to say that atmospheric methyl bromide is 50 times more destructive than CFCs is incorrect and misleading, especially when focusing on anthropogenic effects. Once in the stratosphere, a bromine atom does destroy about 50 times as many ozone molecules as does a chlorine atom.

Atmospheric scientists understand this

distinction but it may help many readers and non-scientist representatives to the Montreal Protocol to clarify this point.

Discoveries of the past several years demonstrate that removal of atmospheric methyl bromide is more rapid than had been thought, and imply the existence of unidentified sources. Are other human-controlled sources at play or are the missing sources mostly natural? Research on such questions should continue, partly because of the need to gauge how much ozone protection we purchase by banning the substance, and partly because the questions are becoming more intriguing. The spirit and provisions of the original Montreal Protocol demanded continued research and periodic assessments of the implications of new results. I hope that we can continue to exercise these provisions.

Ralph Cicerone

Earth System Science,
University of California,
Irvine, California 92697-3100, USA

1. Spurgeon, D. *Nature* **389**, 219 (1997).
2. Yvon-Lewis, S. A. & Butler, J. H., *Geophys. Res. Lett.* **24**, 1227–1230 (1997).

Surgical leadership

Sir — In his review of Stephen S. Hall's *A Commotion in the Blood: Life and Death in the Immune System* (*Nature* **388**, 841; 1997), Fred S. Rosen makes no attempt to conceal his bias about surgeons: "Coley was at heart a surgeon, a profession with a strong theatrical element that is basically in conflict with the painstaking details of data-gathering requisite for good science."

Here it is Rosen who is being theatrical because he lacks supporting data. Surgeons continue to make major scientific advances which are analysed with the same scrutiny as Rosen's own, earning not only peer-reviewed funding but also international recognition.

Perhaps Rosen is unaware that Joseph E. Murray, Charles B. Huggins, Emil T. Kocher and other Nobel prizewinners were all clinically active surgeons.

Rosen lists his affiliation as the Center for Blood Research in Boston. He ought to know that one of the intellectual giants on whose shoulders he stands was Dr Charles S. Drew. Drew, an African-American scientist who must have excelled at the painstaking details of data-gathering to father modern blood-banking, was first and foremost a surgeon.

Surgical leadership abounds in current science despite the pervasive bias that Rosen preaches.

Timothy G. Buchman

Department of Surgery,
Washington University School of Medicine,
660 South Euclid Avenue,
St Louis, Missouri 63110, USA

The ocean's productive deserts

Scott C. Doney

The loss of photosynthetically generated carbon from the ocean's surface waters is a central mechanism in the global carbon cycle. Analyses of the results from three independent techniques give us the best estimate yet of the scale of this 'export' process.

The vast, nutrient-depleted subtropical ocean gyres were for a long period thought to be the marine equivalent of the deserts on land, with little biomass, low overall productivity and a limited role in the carbon cycle. That picture has changed dramatically in the past decade, as highlighted by the work of Emerson *et al.* reported on page 951 of this issue¹.

The authors present three independent field estimates of the annual average carbon flux leaving the subtropical surface ocean at a site in the North Pacific. The different techniques — mass balances of organic carbon itself, and the proxies of dissolved oxygen and inorganic carbon (dissolved carbon dioxide and its products) — converge on a relatively high flux value of about $2 \text{ mol C m}^{-2} \text{ yr}^{-1}$; moreover, Emerson *et al.* show, for perhaps the first time, that we can systematically close the carbon budget for the upper ocean, albeit within large reported error bars ($\pm 50\%$).

Oceanographers have known for some time that the biogeochemistry of the ocean (Fig. 1) is to a large extent driven by the downward flux of organic carbon, and to a lesser extent of inorganic carbon in the form of calcified shells, from the shallow, sunlit surface layer or euphotic zone². Photosynthesis in the ocean is carried out almost entirely by small, floating plants or phytoplankton. Unlike the situation on land, the bulk of this photosynthetically fixed carbon is rapidly recycled either through grazing by zooplankton or by microbial processes, and only a small fraction (somewhere between 10% and 30% in the open ocean) escapes from the surface layer either as sinking particles or as dissolved material transported with the water.

This so-called 'biological pump' controls the partitioning of dissolved inorganic carbon between the surface and deep ocean. It also reduces the partial pressure of carbon dioxide in the surface waters, and thus helps to govern the concentration of atmospheric CO_2 on timescales of millennia. But actually measuring these fluxes has remained an elusive and troublesome task — previous attempts to balance the carbon export with the observed surface drawdown of inorganic carbon by phytoplankton have often failed³.

The traditional approach for measuring

the export flux of organic matter has relied on shallow sediment traps, which are essentially large rain gauges suspended below the base of the euphotic zone to collect any sinking material⁴. As well as the normal issues confronting any environmental sampling, sediment traps in the upper ocean have been plagued with a variety of methodological questions ranging from animal consumption of the trapped material to biases created by hydrodynamic effects in the large, near-surface currents⁵. Shallow traps are typically deployed for only a few days and thus may miss much of the carbon flux if it occurs as episodic events or blooms. Recent work also shows that a significant fraction of the export of organic carbon occurs not as sinking particles but as dissolved organic matter⁶.

Indirect geochemical methods were pioneered by Jenkins and Goldman⁷ to circumvent many of these problems. When phytoplankton produce organic matter through photosynthesis, they take up a corresponding quantity of carbon dioxide from sea water and release oxygen. The reverse reactions occur during respiration, and over seasonal-to-annual timescales the export of organic matter from the surface layer must be balanced by changes in the inventory of dissolved gases. The gas concentrations, however, are also strongly influenced by

exchange with the atmosphere, and vertical mixing and transport by the large-scale ocean circulation, and the crux of the geochemical estimates involves constraining these physical terms so that the net biological production can be calculated by difference.

Emerson and colleagues' analysis¹ is based on data from the Joint Global Ocean Flux Study's Hawaiian Ocean Time-series (HOT) programme, which since 1988 has taken batteries of measurements at the same place at monthly intervals. The annual mean export of organic carbon is computed from particle fluxes into sediment traps, inferred subsurface mixing rates applied to the vertical gradient of dissolved organic carbon, and a minor term due to the net carbon transport by zooplankton, which undergo a vertical diurnal migration. Interestingly, the organic carbon flux is about equally split between particulate and dissolved phases.

Estimates for the oxygen and dissolved inorganic carbon budgets, the proxies for organic carbon, are derived separately. In the case of oxygen, the air-sea gas exchange and vertical mixing rates are found by solving for the vertical profiles of oxygen and the two abiological analogues, nitrogen and argon gas; a similar approach is followed for inorganic carbon using additional information from the distribution of the inorganic ^{13}C isotope. The required vertical mixing rates for both the carbon and oxygen budgets are larger than expected from *in situ* turbulence measurements but could reflect larger-scale circulation effects (see below). The three methods yield values of $2.7 \text{ mol C m}^{-2} \text{ yr}^{-1}$ for oxygen, $2.0 \text{ mol C m}^{-2} \text{ yr}^{-1}$ for organic carbon and $1.6 \text{ mol C m}^{-2} \text{ yr}^{-1}$ for inorganic carbon, but they overlap within the uncertainties of the individual techniques.

But where do the nutrients come from to support such large fluxes of organic carbon? The elemental composition of phytoplankton and marine organic matter generally falls

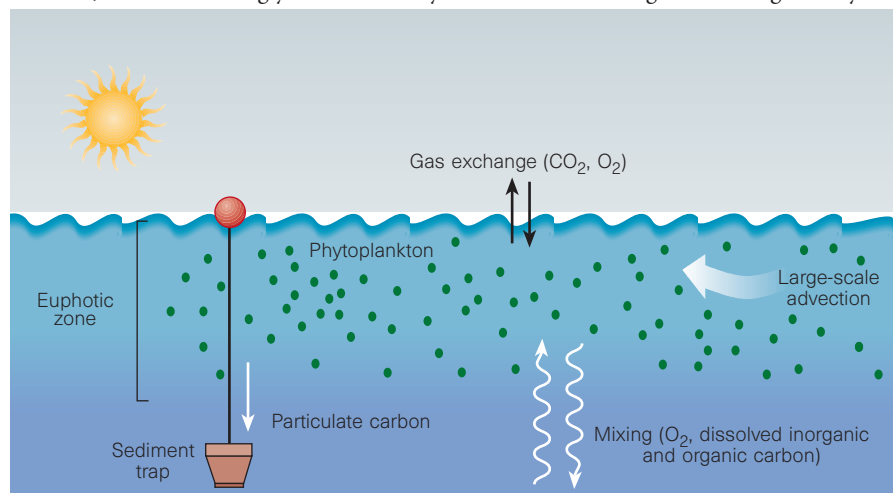


Figure 1 Simplified scheme of biogeochemical cycles in the upper ocean. In the work discussed here, Emerson *et al.*¹ have estimated the downward flux of organic carbon directly from sediment traps and indirectly from the mass balances of oxygen and inorganic carbon (dissolved carbon dioxide and its products). These results bring us closer to closing the carbon budget of the upper ocean.

into a relatively narrow band, and the inorganic nitrogen and phosphorus concentrations in the euphotic zone near the HOT site are generally quite small (at or below the levels that can be detected by traditional methods). Dissolved organic nitrogen and phosphorus compounds⁸ offer one potential supply, but a mechanism is still required to replenish these standing stocks from external sources. Tracer-based measurements⁹ in the thermocline show that small-scale vertical diffusion is much too weak to support the large implied fluxes either of nutrients or of oxygen and carbon.

Some phytoplankton can regulate their buoyancy and thus can mine nutrients, particularly phosphorus, from below the euphotic zone, but this may not be sufficient explanation. Karl *et al.*⁸ have also proposed that substantial levels of nitrogen fixation occur in the subtropical North Pacific, enough to balance about half of the nitrogen export. An alternative hypothesis, offered by McGillicuddy and Robinson¹⁰ among others, is that nutrient inputs are dominated by mesoscale circulation effects (that is, at the level of eddy systems, some several tens to hundreds of kilometres in extent) which could also account for the large predicted vertical fluxes of oxygen, and inorganic and organic carbon¹.

If Emerson and colleagues' findings¹ are representative of other subtropical regions, it seems that these nutrient-poor ecosystems account for an astonishing half of the export of organic carbon from the world's oceans. However, at the only other well-characterized subtropical site (near Bermuda), Michaels *et al.*³, using almost identical analytical methods, find a large carbon imbalance, with the export of organic carbon equalling only about half the inferred uptake of carbon dioxide by phytoplankton. Clearly, an important next step at both sites is to incorporate nutrients, including trace metals such as iron, into a more comprehensive multi-element analysis. But unequivocal closure of the biogeochemical cycles in the subtropical ocean will undoubtedly require better treatment of the mixing and flow of water in the area around the time-series sites, both through physical observations and numerical models.

Finally, at a fundamental level, it is simply not possible to answer the types of questions addressed by Emerson *et al.*¹ and Michaels *et al.*³ from sampling on one or two oceanographic cruises. We need not only more data, but the context that is derived from monitoring a single site; for that we must press the case for continuing support for the taking of long-term, time-series measurements of the oceans. □

Scott C. Doney is at the National Center for Atmospheric Research, Boulder, Colorado 80307-3000, USA.
e-mail: doney@ucar.edu

1. Emerson, S. *et al.* *Nature* **389**, 951–954 (1997).
2. Broecker, W. S. & Peng, T.-H. *Tracers in the Sea* (Eldigio, Columbia Univ., Palisades, NY, 1982).
3. Michaels, A. F., Bates, N. R., Buesseler, K. O., Carlson, C. A. & Knap, A. H. *Nature* **372**, 537–540 (1994).
4. Martin, J. H., Knauer, G. A., Karl, D. M. & Broenkow, W. W. *Deep-Sea Res.* **34**, 267–285 (1987).
5. Gardner, W. D. *Rep. JGOFS Symp.* <http://www.ocean.tamu.edu/JGOFS/contents.html> (1995).
6. Carlson, C. A., Ducklow, H. W. & Michaels, A. F. *Nature* **371**, 405–408 (1994).
7. Jenkins, W. J. & Goldman, J. C. *J. Mar. Res.* **43**, 465–491 (1985).
8. Karl, D. *et al.* *Nature* **388**, 533–538 (1997).
9. Ledwell, J. R., Watson, A. J. & Law, C. S. *Nature* **364**, 701–703 (1993).
10. McGillicuddy, D. J. & Robinson, A. R. *Deep-Sea Res.* (in the press).

Life histories

Be social, live longer

Ross H. Crozier

Ageing — gracefully or not — is something that we humans all endure. With modern medical care and by better diet, exercise and other stratagems, we can extend our active and absolute lifespan to a degree that most would accept as recompense for desisting from gluttony and soft living. But can really large increases be made in lifespans? A report by Keller and Genoud on page 958 of this issue¹ shows that, indeed, they can.

One viewpoint stresses that there are basic cellular and mechanical limits to the body's ability to repair itself — for example, the lenses in our eyes deteriorate with age, and it would be difficult to change the repair schedule². It seems hard to fault this view with regard to the basic question as to why ageing occurs at all. Ageing — an increase in the rate of mortality with age — is expected in species that undergo a division of cells between germ line and soma. Those lineages that skimp on repair mechanisms in favour of reproduction have outcompeted the lineages that have good repair mechanisms (which could maintain the

individuals indefinitely). Although ageless multicellular organisms are known³, there are no immortal ones — even a potentially immortal sea anemone can be stepped on. So, ageing may be the price we pay for complexity.

Given that ageing occurs, how is its rate determined? Since the time of Alfred Russel Wallace, evolutionists have argued that longevity is as capable of evolving as other characteristics. The chief factors are the relationship between age and fecundity, and the risk of mortality. If mortality rates are high, then selection to maintain anatomical and metabolic function in older animals is weaker than if the mortality rates are low — because a smaller proportion of the population is old^{4,5}. Indeed, in laboratory insect populations, longevity can be significantly increased by using the oldest individuals as parents for the next generation. Conversely, there is a considerable decrease in longevity if the youngest individuals are used as parents for the next generation.

A striking case in which ecological conditions mediate longevity comes from studies



Figure 1 Living longer through social security — in their study of social insects, Keller and Genoud¹ have shown that in colonies where one egg-laying queen predominates, she will have a long lifespan. This is because, after a long colony-founding stage, she will face little threat. By contrast, colonies with many queens move often, so the queens will encounter many predators and they have shorter lifespans.

B. HOLDOBLER/THE ANTS, HUP

over several years of abnormal abdomens, due to a ribosomal gene defect, in Hawaiian populations of *Drosophila mercatorum*. The effects of this defect in females include precocious reproductive maturity and decreased longevity. Natural climatic fluctuations which change the probability that the insects will die through desiccation are associated with the expected changes in the gene frequency — drier years lead to an increased frequency⁶.

Laboratory selection for longevity will, at most, double natural lifespans. So might this indicate that there are natural physiological constraints for each kind of animal? Perhaps a fly that is descended from *Drosophila melanogaster*, but which has ten times its longevity, is impossible. Or maybe the amount of evolution that one can cause during one or two grant-lifetimes is insufficient. For example, indicative of the long-term stability of ageing rates, flightless birds (which are subject to higher predation risks) have unusually short lifespans⁴.

Keller and Genoud¹ now provide dramatic support for the evolutionary lability of lifespan in a study of social insects (ants and termites). They tested for an association between longevity and reduced risk of death, using comparative analysis^{7,8}. This is a statistical technique that counts the numbers of evolutionary transitions between characteristics, rather than just counting the number of species. Whereas the workers of social insects do not tend to have especially long lives because of mortality once they become foragers⁹, reproductives in many species can have lifespans of up to decades¹⁰. Keller and Genoud point out that lifespans extended to about ten times ancestral longevity have arisen three times in insects. These increases are associated with the rise of eusociality — a kind of elaborate sociality that is characterized by a reproductive division of labour into egg-laying specialist queens and attendant helper workers. The extended lifespans are plausibly attributed to the need for queens to survive a long colony-founding stage, after which they face very low levels of threat.

Eusocial insects have a range of life-patterns¹⁰. In many species, a single mated queen (in termites, a king and a queen) flies far from her home nest and, in so doing, encounters the perils of numerous predators. She must then endure the privations of a founding period nurturing her first workers. Finally, often after several years, the first descendant reproductives will mature to renew the cycle. In other species, colonies multiply by fragmentation. New queens join established colonies — reducing both the length and danger of the founding period — but these colonies move often, compromising the safety of the queens. So, single-queen colonies should promote extreme longevity, whereas colonies with many queens give a

relatively high pay-off to early reproduction, and less to long life.

Keller and Genoud show that, within ants, extreme longevity is indeed significantly associated with single-queen species. Clearly, there have been no insuperable constraints on the modification of life-pattern in insects. Unfortunately, however, these results give little hope for a dramatic increase in the human lifespan. Such an intrinsic increase would be expected only in the remote future, and the methods to obtain it through selection — such as the universal deferment of reproduction to increasingly later ages — are unlikely to appeal. So is there hope that the key to an indefinite lifespan could be discovered by the isolation of an 'ageing gene'? Alas, the prediction from evolutionary theory is that numerous aspects of the phenotype, and many genes, are probably implicated in limiting life-

span¹¹, although these may be concentrated in a small number of metabolic systems¹². □
Ross H. Crozier is in the School of Genetics, La Trobe University, Bundoora, Victoria 3083, Australia.

e-mail: GENRHC@gen.latrobe.edu.au

1. Keller, L. & Genoud, M. *Nature* **389**, 958–960 (1997).
2. Holliday, R. *News Physiol. Sci.* **7**, 38–40 (1992).
3. Bell, G. *Am. Nat.* **124**, 600–603 (1984).
4. Rose, M. R. *Evolutionary Biology of Ageing* (Oxford Univ. Press, New York, 1991).
5. Charlesworth, B. & Hughes, K. A. *Proc. Natl Acad. Sci. USA* **93**, 6140–6145 (1996).
6. Templeton, A. R., Hollocher, H. & Johnston, J. S. *Genetics* **134**, 475–485 (1993).
7. Harvey, P. H. & Pagel, M. D. *The Comparative Method in Evolutionary Biology* (Oxford Univ. Press, 1991).
8. Felsenstein, J. *Am. Nat.* **125**, 1–15 (1985).
9. Gordon, D. M. & Hölldobler, B. *Psyche* **94**, 341–346 (1987).
10. Wilson, E. O. *The Insect Societies* (Harvard Univ. Press, 1971).
11. Dawkins, R. *River Out of Eden* (Weidenfeld & Nicholson, London, 1995).
12. Ying, W. *Med. Hypoth.* **48**, 143–148 (1997).

Materials science

Invisible circuits

Gordon Thomas

It takes only a little thought to imagine the value of an invisible electronic circuit in both military and consumer applications. Rudimentary circuits, using invisible metals, are already widely used, for example on the digital wrist-watch that I am wearing as I write this article and the flat-panel display screen of the laptop computer I am using. But much more sophisticated invisible circuits are possible, and physicists, chemists and engineers are vigorously pursuing new and improved materials for such circuits, as well as an understanding of their properties. One new material is presented on page 939 of this issue¹, where Hiroshi Kawazoe and his colleagues at the Tokyo Institute of Technology describe an invisible material in which positive charges carry an electrical current. Crucially, p-type materials are needed, along with the more usual n-type materials in

which electrons are the carriers, to make active circuit components such as diodes.

First, let us attack the reasonable misconception that electrical conductors must be visible, as are common metals like copper. The trick to making an electrical conductor that is hard to see with the naked eye (which is what I mean by invisible) is to concoct a material with two special features: a wide enough energy gap between the valence and conduction bands of electronic states; and some, but not too many, electrical carriers.

As shown in Fig. 1, if the energy difference between a full valence band and an empty conduction band is bigger than the energy of a photon, no valence electron can absorb it. So if the gap is bigger than about 3.1 eV (the energy of a blue photon with a wavelength of 400 nm, at the high end of the visible spectrum), the material will be invisible.

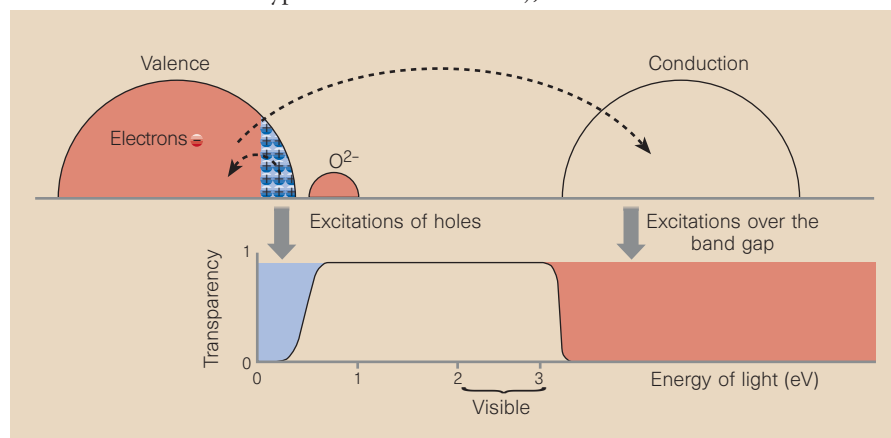


Figure 1 An electronic bandgap designed for transparency. Photons of visible light, between 2.1 and 3.1 eV, do not have enough energy to excite an electron from the valence band to the conduction band; and they have too much energy to excite a hole.

The first challenge, then, is to make and understand materials with bandgaps above 3.1 eV — much larger than the gaps that make conventional electronic materials opaque (around 1 eV for silicon). Towards this end, Kawazoe *et al.* have studied CuAlO_2 and found that it has a suitable bandgap of 3.5 eV. Based on theoretical chemistry, its atoms should have nicely matched valences, with Cu^+ , Al^{3+} and two O^{2-} combining to bind all the charges and form an insulator.

The second challenge is to turn this insulator into an invisible conductor. We can use the insulator as a host for other atoms, and with these impurities we can introduce charge carriers that turn it into a conductor, while leaving its bandgap open. Kawazoe *et al.* grew thin films of CuAlO_2 in an O_2 atmosphere, found that the films conducted electricity, and surmised that some extra oxygen was incorporated. If the extra oxygen atoms form O^{2-} ions, perhaps on interstitial sites, they will bind electrons, leaving behind empty states in the valence band which would act as positive charges (holes), as shown in Fig. 1. The material's chemical formula would then be CuAlO_{2+x} . Measurements of the Hall effect indicated that the carriers indeed acted like holes, and that x was only about 1/50,000.

These added carriers will both absorb and reflect light, and could, in sufficient numbers, make the material quite visible. Common metals, such as copper, are highly reflective of visible light, mostly because its electromagnetic fields induce currents near their surfaces — but only for light up to a characteristic energy, corresponding to the 'plasma frequency'. Because the plasma frequency changes as the square root of the carrier density, our material can remain invisible provided we do not introduce too many carriers. Fortunately, the CuAlO_{2+x} made by Kawazoe *et al.* isn't even close to reflecting visible light strongly — it has so few carriers that its plasma frequency is about 20 times lower than that of red light and about 1,000 times lower than that of copper.

In contrast to CuAlO_{2+x} , nearly all oxides that are good conductors have negative carriers (electrons)^{2,3}. This material is an incompletely understood and very interesting exception. It has been known to exist for 42 years⁴ and to have positive carriers for over a decade⁵, but this is the first time it has been made in a thin film with a d.c. electrical conductivity almost high enough for possible applications. We still do not know why it has been so difficult to make invisible oxides in which the carriers are positive.

A related question is what other materials might also be used in invisible circuits. Besides other oxides, the discussion above applies to the nitrides⁶, such as GaN. The nitrides have been made p-type and n-type, and the two have been made into diodes and even into emitters of blue light⁷. As with the

oxides, it has proved more difficult to make nitrides with positive carriers than with negative ones⁶. The most widely used materials remain the oxides, because the n-type oxides have the highest d.c. electrical conductivities of invisible conductors², more than $1,000 \text{ S cm}^{-1}$.

The material studied by Kawazoe *et al.*¹ has properties characteristic of a doped semiconductor, with a d.c. electrical conductivity of only around 1 S cm^{-1} . As soon as we think of an invisible conductor as a semiconductor, we have made an important step towards imagining new sorts of invisible circuits. In principle, even the most advanced electronics now made with semiconductors can, in principle, be made invisible. In practice, better materials and better understand-

ing are needed. But soon we will be able to go beyond the passive transparent conductors that are components in the flat-panel displays of our watches and computers, to active circuits — including switches, amplifiers, electronic memories, blue lasers and devices not yet invented. □

Gordon Thomas is at Bell Laboratories, Lucent Technologies, 700 Mountain Avenue, Murray Hill, New Jersey 07974, USA.

1. Kawazoe, H. *et al.* *Nature* **389**, 939–942 (1997).
2. Haacke, G. *Ann. Rev. Mater. Sci.* **7**, 73–93 (1977).
3. Phillips, J. M. *et al.* *Appl. Phys. Lett.* **63**, 3335–3337 (1995).
4. Hahn, H. & Lorent, C. Z. *Anorg. Allg. Chem.* **279**, 281–288 (1955).
5. Benko, F. A. & Koffyberg, F. P. *J. Phys. Chem. Solids* **45**, 57–59 (1984).
6. Strite, S. & Morkoc, H. *J. Vac. Sci. Technol. B* **10**, 1237 (1992).
7. Mohammad, S. N. *et al.* *Proc. IEEE* **83**, 1306 (1995).

Signal transduction

Rhomatic interludes raise blood pressure

A. P. Somlyo

In the beginning there was calcium, prime activator of smooth and striated muscle contraction. But the beginning is not the end, as illustrated by the report of Narumiya and colleagues¹ on page 990 of this issue. They describe a drug that inhibits the activity of an enzyme that sensitizes smooth-muscle contraction to Ca^{2+} . Contraction of smooth-muscle cells determines the size of the lumen in blood vessels, airways, the gastrointestinal tract, uterus and bladder, and abnormal contraction can cause diseases such as high blood pressure and asthma. Accordingly, the new drug reduces high blood pressure in experimental animals. The results also throw light on a ubiquitous signalling system that operates not only in smooth muscle, but in most — perhaps all — non-muscle cells. Moreover, the new work raises the possibility of finding causes of, and treatments for, diseases due to abnormalities of smooth-muscle contraction.

When striated muscles are stimulated, levels of cytosolic Ca^{2+} increase and bind to troponin, which is a regulatory protein on thin (actin) filaments. Ca^{2+} binding allows the thin filaments to slide, propelled by crossbridges projecting from adjacent thick (myosin) filaments. This is a tight, relatively direct coupling mechanism between Ca^{2+} and contraction.

In contrast, contraction of vertebrate smooth muscles is activated by a less direct mechanism — Ca^{2+} binds to calmodulin (a cytosolic regulatory protein), and the complex activates myosin light-chain kinase (Fig. 1). This enzyme phosphorylates a light chain (small subunit) of myosin, resulting in contraction, which also occurs by a sliding-filament mechanism^{2,3}. Relaxation is usually

initiated by a fall in levels of Ca^{2+} . It is mediated through dephosphorylation of the myosin light chain by a protein phosphatase which is directed, by its regulatory subunit, to target myosin³. This indirect, 'loose coupling' to Ca^{2+} allows contraction to be regulated by mechanisms that, at a constant concentration of Ca^{2+} , increase or decrease the activities of myosin phosphatase^{4,5} or light-chain kinase⁶. A decrease in the activity of myosin phosphatase is referred to as Ca^{2+} sensitization, and an increase in the activity of this enzyme (or a decrease in activity of the kinase) is referred to as Ca^{2+} desensitization.

The story began 'downstream' when, as predicted⁴, activation of certain guanosine-nucleotide-binding (G) proteins was found⁵ to inhibit myosin phosphatase and, hence, to increase myosin light-chain phosphorylation and contraction, independently of Ca^{2+} . The Rho isoform, RhoA, belongs to the Ras superfamily of small (relative molecular mass ~20K) G proteins that are inactive when complexed with GDP, but active in the GTP-bound form. Members of the Ras family are involved in many cell functions and dysfunctions, including cellular trafficking, enzyme regulation and cancer. RhoA is an important 'upstream' messenger of Ca^{2+} sensitization. When exogenous, constitutively active RhoA is added to permeabilized smooth muscles, myosin phosphorylation and force are increased at constant concentrations of Ca^{2+} . In contrast, inactivation of cellular RhoA by bacterial enzymes inhibits the Ca^{2+} -sensitizing effects of hormones, drugs and other agents that activate RhoA^{7,8}.

In resting smooth muscle, RhoA is in an inactive, GDP-bound form. Most of it is in the cytosol where it is complexed with another

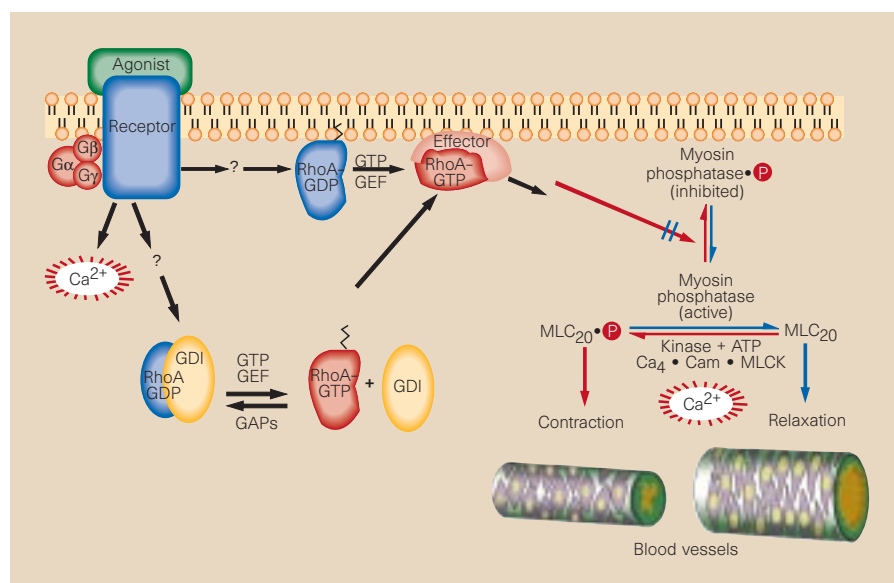


Figure 1 Narumiya and colleagues¹ have identified a drug (Y-27632) that inhibits the activity of a Ca^{2+} -sensitizing enzyme, leading to a reduction of high blood pressure in experimental animals. Activation of receptors coupled to certain guanine-nucleotide-binding (G) proteins releases intracellular Ca^{2+} that binds to calmodulin (Cam), and this complex activates myosin light-chain kinase (MLCK). By phosphorylating the regulatory light chain of myosin (MLC_{20}) in smooth muscle, MLCK causes vascular smooth muscle to contract and the lumen of blood vessels to narrow. Many of the same receptors also activate RhoA and, with the help of guanine-nucleotide exchange factors (GEFs), dissociate cytosolic RhoA-GDP from guanine-nucleotide dissociation inhibitor (GDI); this allows the exchange of GTP for GDP on RhoA. The active RhoA-GTP activates Rho-associated kinase which phosphorylates — and so inhibits — myosin phosphatase. Myosin phosphatase dephosphorylates smooth-muscle myosin, causing the smooth muscle to relax and blood vessels to dilate. Y-27632 inhibits Rho-associated kinases, thereby blocking the inhibition of smooth-muscle myosin phosphatase and Ca^{2+} sensitization. So, although Ca^{2+} is the main activator of smooth-muscle contraction (through MLCK), the level of force can be modulated independently of it. (Figure adapted from ref. 18; Jama Courtney helped in its preparation.)

er protein called GDP dissociation inhibitor, although a smaller amount of inactive RhoA is also associated with the cell membrane. Like other G proteins, RhoA becomes active when, with the help of exchange factor(s) coupled to membrane receptors in an as-yet unknown manner, it replaces GDP with GTP. Cytosolic RhoA-GTP, dissociated from the GDP dissociation inhibitor, is translocated to the plasma membrane⁹.

This creates a paradox, because the myosin phosphatase that RhoA inhibits is in the cell interior. The 'matchmaker' between membrane-associated RhoA and cytosolic myosin phosphatase is another enzyme, or family of enzymes — Rho-associated kinases. When activated by RhoA, these enzymes phosphorylate the regulatory subunit and so inhibit the catalytic activity of smooth-muscle myosin phosphatase^{10–12}. The result is Ca^{2+} sensitization: increased myosin phosphorylation and contraction at constant concentrations of Ca^{2+} (ref. 13).

Increased contraction narrows the lumen of blood vessels and increases the resistance to blood flow, resulting in high blood pressure. Narumiya and colleagues¹ have produced a relatively specific and high-affinity inhibitor of Rho-associated kinases. This drug (code-named Y-27632) not only

inhibits Ca^{2+} sensitization of vascular smooth muscle, but also reduces high blood pressure in animals. This suggests that overactivity of the RhoA pathway, rather than (as previously believed) increased Ca^{2+} alone, can cause high blood pressure. This is a considerable step towards determining the pathophysiological roles of RhoA and its effectors, the more remarkable because the drug has only a minimal effect on normal blood pressure.

The story goes further, because RhoA is but one of several members of the Rho subfamily of Ras proteins. The amino-acid sequences and crystal structures of proteins that belong to this subfamily, including CDC42 and Rac, are subtly different from one another^{14,15}. They interact with different effectors that are involved in multiplex, often crosstalking, signalling. For example, active RhoA — but not Rac — participates in transcriptional regulation and cell division¹⁶, and induces the formation of stress fibres, which are bundles of actin filaments interspersed with myosin and α -actinin in cultured cells.

But only some of the effects of RhoA are mediated by Rho-associated kinase(s). Stress-fibre formation has been attributed to a mechanism that is analogous to Ca^{2+} sensitization in smooth muscle — inhibition of (non-muscle) myosin phosphatase¹⁷. The Narumiya



100 YEARS AGO

A suggestive paper by Prof. W. P. Mason, entitled "Sanitary Problems connected with Municipal Water Supply," has been published in the *Journal of the Franklin Institute*. ... The writer tells us that the average annual typhoid death rates for thirteen Massachusetts cities before the introduction of a public water supply was 7.94 per 10,000, whilst since the improvements have been carried out the deaths from typhoid fever have fallen to 3.83 per 10,000. In the whole State of Connecticut the percentage of typhoid deaths to total deaths has fallen from about 5.8 in 1870 to 1.84 in 1893. ... A very remarkable example of how typhoid fever may be spread is given in the case of a serious outbreak of this disease which took place at Plymouth, Pa. The origin of this disastrous epidemic was traced to a single typhoid patient whose dejecta were thrown out upon the snow of a frozen hillside, at the base of which ran a small stream, whence the town water supply was ultimately drawn. Several weeks elapsed, during which the dejecta were hard frozen before the March thaws permitted the melting snows to wash them into the stream below; but during this interval the typhoid germs had retained their vitality and full complement of virulence, as demonstrated by the otherwise quite unaccountable outbreak of typhoid fever in the said town.

From *Nature* 28 October 1897.

50 YEARS AGO

After the introduction of newer synthetic antimalarials, the possibility of acquired resistance of plasmodium strains to these drugs became important, analytically and clinically. ... In my series of experiments I started with quinine; then I tried to produce resistance to 'Paludrine', and I can confirm the possibility of resistance. Among the strains in my laboratory there is one (*Plasmodium gallinaceum*) which by administering seven times 4 mgm./kgm. 'Paludrine' weekly had acquired after some fifteen weeks a high degree of resistance to 'Paludrine'. In contrast with Williamson, Bertram and Lourie, I was able to produce in *Plasmodium gallinaceum* a moderate, nevertheless clearly demonstrable, resistance to quinine. ...

From *Nature* 1 November 1947.

group¹ now shows that their Rho-associated kinase inhibitor also inhibits the formation of stress fibres. On the other hand, this compound does not affect the transcriptional effect of RhoA, nor does it inhibit the activity of p21-activated serine/threonine kinase, which is activated by CDC42. Therefore, Y-27632 should be useful in sorting out, from among the many crosstalking, G-protein-coupled pathways, those that are mediated by Rho-associated kinases. It remains to be seen whether high blood pressure in humans, a major cause of heart disease and stroke, is due to disorders of RhoA and Rho-associated kinases and treatable by their inhibitors. But the new findings should certainly draw attention to Ca^{2+} -independent causes of, and treatments for, diseases due to abnormalities of smooth-muscle contraction. □

A. P. Somlyo is in the Departments of Molecular Physiology and Biological Physics and Internal Medicine (Cardiology), and the Center for Structural Biology of the University of Virginia, PO Box 10011, Charlottesville, Virginia 22908, USA.

Athletics

Mostly in the mind

Tim Lincoln

Two athletes in a light aircraft wander way off course, become utterly lost and crash-land. Both are unharmed and one sets off to ask a person in the distance where they are. "47° 30' North, 19° 05' East," comes the reply. "Well," says one of the athletes to the other, on hearing this answer, "we still don't *really* know where we are, but we do know that was a sports physiologist."

This story was recounted at a meeting this month* which brought together scientists, coaches and athletes to ruminate about the potential for improvement in records in track and field events. At one end of the spectrum there were volleys of graphs and histograms from some of the scientists — hence the self-deprecating story, told by the physiologist P. E. Di Prampero (Univ. Udine, Italy), acknowledging the difficulty in translating detailed scientific information and theories into helpful advice for athletes.

At the other, there were the personal reflections of such performers as Sergey Bubka, holder of the pole-vault world record (at 6.14 m), and of Ana Fidelia Quirot, who recovered from horrifying burn injuries to win the women's world championship 800 metres in 1995. Kip Keino, Olympic gold-medallist in 1968 and 1972, and former world-record holder at middle distance, weighed in with an account of the reasons behind the rise and sustained success of Kenyan runners at world level.

1. Uehata, M. et al. *Nature* **389**, 990–994 (1997).
2. Hartshorne, D. J. in *Physiology of the Gastrointestinal Tract* 2nd edn (ed. Johnson, L. R.) 423–482 (Raven, New York, 1987).
3. Somlyo, A. P. & Somlyo, A. V. *Nature* **372**, 231–236 (1994).
4. Somlyo, A. P. et al. *Adv. Prot. Phosphatases* **5**, 181–195 (1989).
5. Kitazawa, T., Masuo, M. & Somlyo, A. P. *Proc. Natl Acad. Sci. USA* **88**, 9307–9310 (1991).
6. Gallagher, P. J., Herrington, B. P. & Stull, J. T. *J. Muscle Res. Cell Motil.* **18**, 1–16 (1997).
7. Otto, B., Steusloff, A., Just, I., Aktories, K. & Pfitzer, G. *J. Physiol. (Lond.)* **496**, 317–329 (1996).
8. Gong, M. C. et al. *Proc. Natl Acad. Sci. USA* **93**, 1340–1345 (1996).
9. Gong, M. C., Fujihara, H., Somlyo, A. V. & Somlyo, A. P. *J. Biol. Chem.* **272**, 10704–10709 (1997).
10. Ichikawa, K., Ito, M. & Hartshorne, D. J. *J. Biol. Chem.* **271**, 4733–4740 (1996).
11. Kimura, K. et al. *Science* **273**, 245–248 (1996).
12. Lim, L., Manser, E., Leung, T. & Hall, C. *Eur. J. Biochem.* **242**, 171–185 (1996).
13. Kureishi, Y. et al. *J. Biol. Chem.* **272**, 12257–12260 (1997).
14. Hirshberg, M., Stockley, R. W., Dodson, G. & Webb, M. R. *Nature Struct. Biol.* **4**, 147–151 (1997).
15. Wei, Y. et al. *Nature Struct. Biol.* **4**, 699 (1997).
16. Hall, A. *Annu. Rev. Cell Biol.* **10**, 31–54 (1994).
17. Burridge, K. & Chrzanowska-Wodnicka, M. *Annu. Rev. Cell Dev. Biol.* **12**, 463–519 (1996).
18. Fujihara, H. et al. *Mol. Biol. Cell* **8**, 2437–2447 (1997).

Di Prampero's work exemplified the theme of the presentations on human biology, which is continued refinement of data on individual athletes that can be fed back into their training regimes. His model takes three main parameters: the energy cost of running per unit of distance, maximum oxygen consumption and maximum amount of energy derived from anaerobic stores. These measurements were carried out on middle-distance runners, predictions made as to their best times, and the results compared with actual performances. Di Prampero says that they agreed well, and the approach has

since been applied to competition cyclists¹: specific training could possibly be used to tweak the parameters to advantage for both athletes and cyclists.

There are also innovative ways of looking at the human body as a machine (J. Diamond, Univ. California Los Angeles; ref. 2). This research has involved calculating sustained metabolic rates of animals in the wild, of experimental mice undergoing intensive heat or milk production (and thus energy use), and of cyclists in the Tour de France. The overall conclusion is that where-as heart, lung and muscle are generally taken to be the organs that determine athletic performance, and so are monitored, the intestine, liver and kidneys are also important for energy budgets averaged over long times. These concepts are not directly applicable to track and field athletes generally, for few events are related to endurance as defined in this metabolic scheme. But again, they could be applied in training — to help devise individual diets that reduce the work done by the intestine, liver and kidneys, or to test the performance of these organs in search of factors that might limit training intensity.

Other areas of scientific contribution are the development of ever-more-refined tracks (G. Gambino, MONDO S.p.A., Italy). That track improvement has come a long way was graphically illustrated by a different speaker, a 400-metre runner from the days of cinder surfaces, when it might be best to draw the outside lane and thus the one least churned up in that part of the track used in the run-up for the javelin.

Consideration of the composition of the track can now be as sophisticated as calculation of potentially adverse heat transfer from track to shoe to the foot of the athlete during events such as the 10,000 metres, when the competitors make some 6,000–7,000 'foot placements' each. More contentiously, a debate polarizing sprinters and longer-



Figure 1 Dorando Pietri half staggers and is half helped across the finishing line of the marathon in the 1908 Olympics — an enduring image of physical and mental toughness in athletics. Pietri's courage was to no avail, because he was disqualified for receiving assistance. He was, however, presented with a special gold cup as a consolation.

*Human Performance in Athletics: Limits and Possibilities, International Athletic Foundation Seminar, Budapest, Hungary, 11–12 October 1997.

distance runners is over the 'hardness' of the running surface: the former favour harder tracks. An answer may lie in the nine-lane stadium, with the inside lane being designed for higher energy absorption and hence for use in long- and middle-distance races. A further advantage of this arrangement, depending on the geometry, could be that the other eight lanes would have a shallower curve, to the particular benefit of the 200- and 400-metre runners.

Shoe design is also evolving. One approach stems from the recognition that the human foot is anatomically less suited for running and jumping than that of some animals; the goal here is to produce a shoe that constrains the foot and ankle to perform like, say, those of a cheetah, and yet is acceptable to athletes — and, presumably, officialdom. Another is to employ finite element analysis in examining the force transfer between shoe and track, with the aim of tailoring the spike plate to provide the most efficient combination of geometry and materials for the best traction for each event (C. Fusco, Adidas, Portland, Oregon).

So far, so positive. But two spectres hang over athletics. One is that, at the top level, its future depends on being successful as a business in fighting for sponsorship, TV coverage and public attention against rival attractions (most notably soccer). The other spectre, drug-enhanced performance and record-breaking, cropped up intermittently — to the dismay of some participants who feel that the malign influence of doping has been overstated. Nonetheless, the entire credibility of athletics, both as a public spectacle and as a human endeavour, hinges on drug eradication or at least its effective policing. Sir Roger Bannister tackled the issue head on. Bannister holds a unique position, being a scientist and an international icon for his achievement, in 1954, in beating the four-minute barrier for the mile (a time now bettered by 16 seconds, just one indication of improved performance over the years).

He is encouraged by the paradoxical and astonishing decline, since the late 1980s, in best distances achieved each year in all throwing events except the javelin. Thanks to some exhaustive detective work involving analysis of Stasi records³, that decline can now be put down unequivocally to the artificial inflation in standards by the state-promoted use of steroids by East German and other countries' athletes before out-of-competition testing was introduced in 1988. Since then, testing has become tougher and the Berlin Wall has come down. The pigeons are still coming home to roost, however, in the continuing emergence of ill health in the subjects treated with steroids.

The doping question is far from straightforward. Legally, governing bodies face crippling legal fees and payment of damages to banned performers who sue them. No

athlete accused of drug use has yet won such a suit outright, but the threat remains. Scientifically, there is a grey area between proscribed and approved drugs, and at present it is difficult to detect banned substances such as the natural hormones human chorionic gonadotropin and erythropoietin (the one stimulates production of testosterone, a steroid, the other increases the number of oxygen-carrying red blood cells). There is also a need for acceptance of the routine use of blood (rather than urine) samples. All in all, it may be through analytical chemistry that science has the most to offer athletics.

There was, however, one ever-present constant for all of the speakers with a first-hand experience of winning or record-breaking — the overwhelming importance of mental attitude in distinguishing the

truly great athletes. As Fusco pointed out, the advantages of quantification of an athlete's day-to-day psychology would be enormous. As it was, athletes and coaches were uncompromising in their refusal to acknowledge physiological limits to performance. For Bubka the aim in the pole vault is now 6.20 metres; for the 100-metre sprinters coached by John Smith (Univ. California Los Angeles) it is a time of 9.76 seconds against a current world record of 9.84. If there was one conclusion to be drawn from the meeting, it was that athletics records will continue to tumble simply because they're there. □

Tim Lincoln is News and Views Editor of Nature.

1. Capelli, C. *et al. Med. Sci. Sports Exerc.* (in the press).
2. Hammond, K. A. & Diamond, J. *Nature* **386**, 457–462 (1997).
3. Franke, W. W. & Berendonk, B. *Clin. Chem.* **43**, 1262–1279 (1997).

Biom mineralization

A pavement of pearl

L. Addadi and S. Weiner

Of all the many organisms that form mineralized hard parts, the molluscs probably most deserve the title of 'master builders'. They form their shells out of calcium carbonate (calcite and/or aragonite) and a matrix of biological macromolecules. These components are arranged in complex architectures, resulting in lightweight materials with high mechanical performance. Shells are composites formed by such superbly controlled processes as to induce the envy and admiration of any materials scientist. New microscopic techniques have now added a piece to the puzzle of how one common shell material, nacre, grows — its matrix sheets have pores that both allow the crystals to grow and keep them aligned¹.

Mollusc shells have at least seven grossly

different types of architecture, among which nacre is the best studied². Nacre is the internal lustrous 'mother of pearl' layer of many common shells (Fig. 1), and pearls themselves are composed entirely of nacre. The structure is a brick and mortar arrangement³: the bricks are flat polygonal crystals of aragonite (Fig. 2), and the mortar is made out of polysaccharide and protein fibres laid down orthogonally to each other and aligned with the aragonite crystal axes⁴.

The cells responsible for shell formation first construct the organic matrix, and then, from a distance, induce the crystals to form within the available spaces. But how do they build a wall, when the mortar is already in place and the bricks must be formed *in situ*? One obvious problem is how to allow for the



Figure 1 **Mother of pearl** — *Neotrigonia margaritacea* shows off its shiny nacre lining.



Figure 2 **The 'stack of coins' structure of aragonite crystals that make up gastropod nacre. This is a scanning electron micrograph of a fractured shell.**

flow of materials (ions, proteins and water) to and from the sites where the crystals are formed. Another problem is how to grow a series of crystals in almost perfect alignment.

Members of the UC Santa Barbara Materials Research Center have now thrown light on both these questions. Tilman Schaffer and colleagues¹ studied demineralized pieces of 'flat pearl'; nacre deposited by the red abalone (a gastropod mollusc) on a glass coverslip introduced between the shell and the adjacent layer of living cells called the mantle. They show that many pores are present in the matrix sheets, through which ions and presumably also proteins can flow. The crystals can thus grow far away from the cells of the mantle, which supply the raw materials for growth.

Schaffer *et al.* used an atomic force microscope, combined with the new technique of scanning ion conductance microscopy. In that technique, a thin tip containing one of the electrodes is vibrated and scanned over the matrix sheets under water. Areas of low and high ion conductance were monitored, proving that the sheets are transparent to ions. The authors also suggest that some of the pores allow crystals to grow from one layer to the next without interruption (Fig. 3). A new nucleation event would thus not be required for every crystal in a stack. The growth through mineral bridges, as suggested, may ensure perfect alignment of the crystal axes along a stack without disrupting the basic brick and mortar alternation which provides the material with elasticity and increased resistance to fracture.

Mollusc nacre has always been at the vanguard of biomineralization studies, probably because of its elusively simple geometry. In 1847 Carpenter described its layered structure³, and in 1924 Schmidt established the presence of discontinuous organic material⁶. By 1930, Bøggild in his classic study of mol-

Galaxy evolution

Smashed into shape



Collisions of galaxies are usually spectacular, but the latest images of a nearby cosmic pile-up do more than impress the eye.

This colour image from the Hubble Space Telescope (B. Whitmore, STScI) shows the impact region of two merging spirals, the Antennae galaxies. The region is outlined on the ground-based black-and-white image.

Star formation, induced by the collision, is clearly traced by the blue light from young, hot stars. Surprisingly, many

compact young clusters of stars are seen in this and other recent observations of colliding galaxies. Previously, stellar groups of this size and shape, such as the globular clusters of our own Galaxy, were thought only to contain much older stars.

The clusters may also be useful as cosmic chronometers: by using the ages of the member stars to tell when a collision took place, it may be possible to investigate how interacting galaxies evolve, perhaps from spiral to elliptical shapes.

Karen Southwell

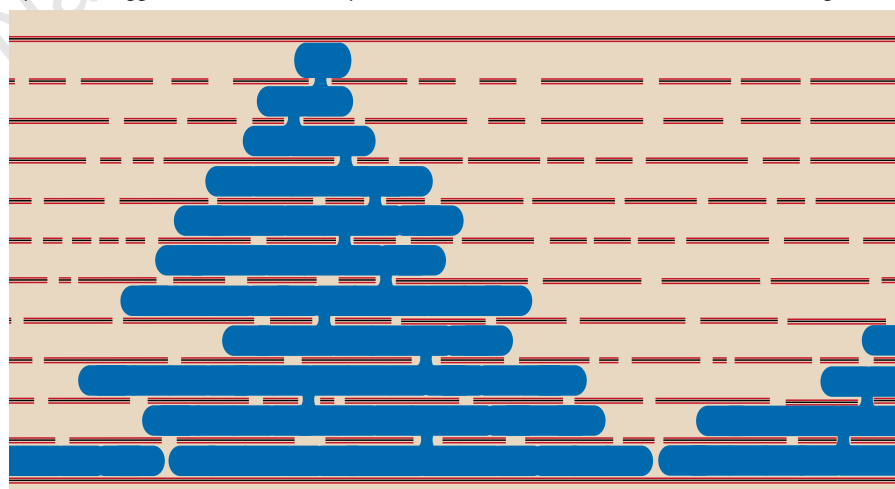


Figure 3 Crystals of mollusc nacre growing between pre-deposited matrix sheets. The sheets have a central layer of core fibres (black), covered by protein layers (red), and are pierced by small pores that allow materials to be supplied to the growing aragonite crystals (blue), as well as keeping the layers crystallographically aligned via mineral bridges. (Adapted from a transmission electron micrograph⁸.)

lusc-shell architectures, described nacre as "so well known, that it may be treated here briefly"². Yet to this day it is the subject of many investigations.

The modern-day pioneer of the study of mollusc-shell formation, Charles Gregoire, discovered holes in the matrix using trans-

mission electron microscopy, and noted that they had different shapes in different species⁷. In the abalone, the observed pore density was approximately $100 \mu\text{m}^{-2}$, with an average diameter of 43–49 nm. The later study of abalone matrix by Nakahara and colleagues revealed a core 10 nm thick that is electron-lucent (relatively transparent to electrons), sandwiched between two 40-nm-thick electron-dense layers⁸.

All these measurements are in remarkable agreement with those of Schaffer *et al.*¹. On the interlamellar sheets they see rounded features separated by pits, with a pit density of $97 \mu\text{m}^{-2}$. Partial digestion with proteolytic enzymes reveals pores 5–50 nm in diameter, and fibre-like structures about 10 nm thick. The fundamental difference from the previous measurements, which makes these data so reliable and convincing, is the use of atomic force microscopy. The specimen is observed unstained and still hydrated, avoiding the well-known artefacts that electron microscopy can introduce.

Schaffer and colleagues are very timely in solving this piece of the puzzle, because other pieces are also falling into place. For example, a biochemical study has shown that an important component of mollusc nacre is a protein similar to silk⁹. Perhaps the largest

remaining area of *terra incognita* is the role of the soluble glycoproteins, which are also important matrix components, located on the matrix surfaces and within the crystals themselves. They are widely thought to play an important part in nucleation and in controlling crystal growth, but are still poorly understood. The first complete amino-acid sequence of one of these proteins was published last year, showing that it has carbonic anhydrase activity (implying that it provides and regulates carbon ions required for aragonite formation), and has structural repeats characteristic of fibrous proteins and proteins that are interacting with surfaces¹⁰.

The field of biomineralization has not yet reached maturity, in the sense that it does not have a proven conceptual framework for understanding the underlying processes.

The studies of mollusc nacre are at the very forefront of this quest — the pearl in the crown, so to speak. □

L. Addadi and S. Weiner are in the Department of Structural Biology, The Weizmann Institute of Science, Rehovot 76100, Israel.

1. Schaffer, T. E. *et al.* *Chem. Mater.* **9**, 1731–1740 (1997).
2. Boggild, O. B. K. *Dan. Vidensk. Selsk. Skr. Naturvidensk. Math. Afd.* **9**, 233–326 (1930).
3. Watabe, N. J. *Ultrastruct. Res.* **12**, 351–370 (1965).
4. Weiner, S. & Traub, W. *Phil. Trans. R. Soc. London Ser. B* **304**, 421–438 (1984).
5. Carpenter, W. *Rep. Br. Ass. Adv. Sci. 17th Meeting*, 93–134 (1847).
6. Schmidt, W. J. *Die Bausteine des Tierkörpers in Polarisiertem Lichte* (Cohen, Bonn, 1924).
7. Grégoire, C., Duchateau, G. & Florin, M. *Ann. Inst. Oceanogr. Belg.* **31**, 1–36 (1955).
8. Nakahara, H., Bevelander, G. & Kakei, M. *Venus, Japan. J. Malacol.* **41**, 33–46 (1982).
9. Sudo, S. *et al.* *Nature* **387**, 563–564 (1997).
10. Miyamoto, H. *et al.* *Proc. Natl. Acad. Sci. USA* **93**, 9657–9660 (1996).

Animal behaviour

Swimming against the current

R. Glenn Northcutt

As any stream-wading child or angler knows, fishes and many amphibians orientate their bodies upstream. This behavioural response to currents — known as rheotaxis — helps young aquatic animals to counteract passive drift. It also allows animals to maintain a position within a stream (station-keeping), maximizes their perception of chemical clues and interception of prey, and minimizes energy expenditure during long-distance migrations. How do these aquatic animals perceive currents? On page 960 of this issue, Montgomery *et al.*¹ provide evidence that one class of lateral-line receptors is involved in rheotaxis.

Initially, it was suggested that aquatic animals detect pressure changes as they are displaced downstream. But experiments at the turn of the century clearly implicated visual and tactile clues². When fishes are placed in an artificial stream or in an experimental apparatus that simulates the bottom of a moving stream, they immediately swim against the current or with the apparent

movement of the bottom. Blinded fishes, however, swim about randomly until they touch the bottom, and then they turn into the current. Some early investigators believed that fish and many amphibians should be able to detect currents with the lateral-line system — a series of mechanoreceptive organs distributed over the head and trunk, and sensitive to relative movement of the surrounding water. But Sven Dijkgraaf, an eminent authority on the lateral-line system, repeated these experiments and concluded that it plays little, or no, role in the detection of currents³.

Dijkgraaf's reputation for meticulous work and profound insight was so great that subsequent reviewers^{4,5} did not even discuss the possibility that the lateral-line system might be involved in rheotaxis. But in the past few years, theoretical⁶ and experimental⁷ studies have indicated that one class of lateral-line receptors should — and do — respond best to the low-frequency, constant-velocity water motions that are most

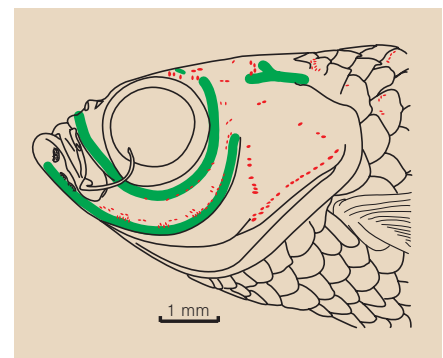


Figure 2 The distribution of the neuromasts of the mechanoreceptive lateral-line system in a zebrafish is typical of that in many fishes. One class of neuromasts is enclosed in canals (indicated by green bands), and these neuromasts are primarily sensitive to accelerating water motions. A second class (indicated by red dots) occurs superficially on the surface of the skin, and these neuromasts are mainly sensitive to low-velocity water movements. Montgomery *et al.*¹ have blocked the function of each set of neuromasts and they find that, in these fishes, the superficial neuromasts are critical for the initiation of rheotaxis (a behavioural response to currents).

likely to cause rheotactic behaviour.

The basic sensory unit (Fig. 1a) in the mechanosensory lateral-line system of fishes and amphibians consists of neuromasts. These are multicellular organs composed of a centrally elongated strip of hair-cell receptors, surrounded by a zone of supporting cells. The apical surface of each hair cell (Fig. 1b) consists of a ramp of short microvilli and an elongated single cilium (kinocilium). Each hair cell is said to be a polarized receptor, because bending of the microvilli towards the kinocilium depolarizes the cell, whereas bending in the opposite direction hyperpolarizes the cell. Most of the hair cells form pairs, the kinocilia of which face in opposite directions (Fig. 1b). Neuromasts are normally sensitive only to displacements that are parallel to the long axis of their sensory strips. The kinocilia are usually embedded in a gelatinous, dome-shaped cupula which acts mechanically to facilitate transduction.

Most fishes contain two classes of neuromasts, which differ in their location, morphology, reaction to pharmacological agents and response properties. The neuromasts of one class are located in canals beneath the skin of the head and trunk; those of the second class occur superficially on the skin of the head and trunk. Canal neuromasts are generally larger, and they possess more hair cells than superficial neuromasts. Their neural activity can be blocked by exogenous exposure to the antibiotic gentamicin⁸, and they respond best to accelerating (or high-frequency) water motions, of the kind generated by a potential predator or prey⁷. In

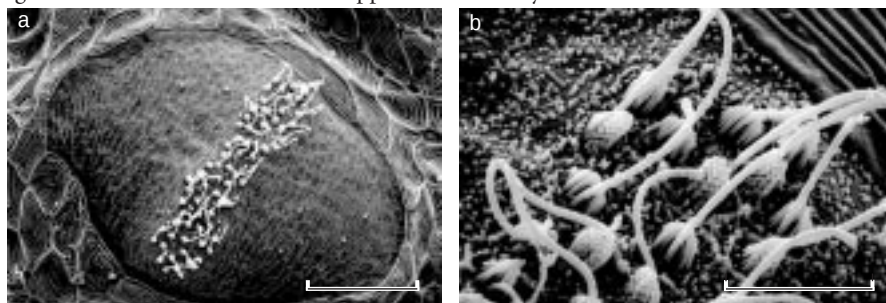


Figure 1 The mechanosensory lateral-line system of fishes and amphibians consists of sensory organs termed neuromasts, and the cranial nerves that innervate these organs. a, Each neuromast consists of an elongated sensory strip of hair cells surrounded by a zone of supporting cells. b, The apical surface of each hair cell consists of a ramp of short microvilli and a single, longer cilium (kinocilium). The overlying cupula has been removed, to allow the superficial surface of the neuromast and its cell types to be seen. Scale bars: a, 20 μ m; b, 5 μ m.

contrast, the neural activity of superficial neuromasts is not affected by gentamicin, and they respond best to low-frequency (or constant-velocity) water motions, such as those that occur in slow-moving streams.

Fortuitously, the neural activity of neuromasts in both classes can be blocked by exogenous exposure to cobalt ions or streptomycin^{9,10}. Montgomery *et al.*¹ used these pharmacological properties to dissect the potential actions of canal and superficial neuromasts in rheotaxis. They found that when the activity of both classes of neuromasts was blocked, fishes could orientate themselves to a water current — but only when the velocity of the current was considerably higher than that for control animals. When they blocked the activity of the canal neuromasts only, however, these fishes showed the same threshold for rheotactic behaviour as control fishes. So superficial neuromasts are critical for the initiation of rheotaxis, and the authors confirmed this by ablating these receptors.

Given these results, it would be simple to infer that superficial neuromasts detect currents and provide the clues needed to initiate rheotaxis, whereas canal neuromasts detect changes in velocity and provide the clues needed to attack prey and avoid predators. In other words, we could say that we now understand the biological significance of superficial and canal neuromasts. But nothing could be further from the truth and, wisely, Montgomery *et al.* do not overstate their conclusions. The lateral-line system of fishes and amphibians shows a staggering range of variation¹¹. All neuromasts are superficial in lampreys and amphibians, although amphibians have homologues of both of the classes of neuromasts that are found in bony fishes. So does the loss of canals in modern amphibians have an effect on neuromast function? The lateral-line

system in bony fishes is equally perplexing. Living species from earlier radiations of bony fishes (bichirs, sturgeons, gars and bowfins) have a limited number of superficial neuromasts. By contrast, many teleosts — the latest group of bony fishes to evolve — have increased the number of superficial neuromasts to such an extent that they greatly outnumber canal neuromasts. It is unlikely that this range of morphological variation, with its functional implications, is the product of selection acting on only two behaviours. Studies are continuing to uncover hitherto unknown properties of these receptors¹², as well as the part that they play in generating the complex cognitive world of aquatic animals. Such discoveries will continue to occupy neuroethologists for generations to come. □

R. Glenn Northcutt is at the Neurobiology Unit, Scripps Institution of Oceanography and the Department of Neurosciences, University of California at San Diego, La Jolla, California 92093-0201, USA.

1. Montgomery, J. C., Baker, C. F. & Carton, A. G. *Nature* **389**, 960–963 (1997).
2. Lyon, E. P. *Am. J. Physiol.* **12**, 149–161 (1904).
3. Dijkgraaf, S. *Biol. Rev.* **38**, 51–105 (1963).
4. Coombs, S., Görner, P. & Münz, H. (eds) *The Mechanosensory Lateral Line* (Springer, New York, 1989).
5. Bleckmann, H. *Prog. Zool.* **41**, 1–115 (1994).
6. Kalmijn, A. J. in *The Mechanosensory Lateral Line* (eds Coombs, S., Görner, P. & Münz, H.) 187–215 (Springer, New York, 1989).
7. Kroese, A. B. A. & Schellart, N. A. M. *Neurophysiology* **68**, 2212–2221 (1992).
8. Song, J., Yan, H. Y. & Popper, A. N. *Hearing Res.* **91**, 63–71 (1995).
9. Karlson, H. E. & Sand, O. J. *Exp. Biol.* **133**, 249–262 (1987).
10. Blaxter, J. H. S. & Fuiman, L. A. in *The Mechanosensory Lateral Line* (eds Coombs, S., Görner, P. & Münz, H.) 481–499 (Springer, New York, 1989).
11. Northcutt, R. G. in *The Mechanosensory Lateral Line* (eds Coombs, S., Görner, P. & Münz, H.) 17–78 (Springer, New York, 1989).
12. Montgomery, J., Coombs, S. & Janssen, J. *Brain Behav. Evol.* **44**, 299–306 (1994).

and Ramakrishnan, colloquially known as the ‘Gang of Four’) introduced scaling concepts to the study of conductance⁴ in such systems. Scaling methods came into prominence some years ago in the theory of phase transitions, and are based on integrating out fluctuations on short length scales to obtain a differential equation describing the evolution of some physical property as the maximum length scale is varied.

Abrahams *et al.* analysed the evolution of the conductance of a block of material as the length of its sides is increased. For a two-dimensional metal, the elementary result is that the conductance is independent of the length of the sides; but Abrahams *et al.* found a small correction to this, caused by back-scattering of the electrons from the weak random impurity potential. This leads to localization of the electrons, but with a large length scale (weak localization).

Two-dimensional systems mark the borderline between high and low dimension, as far as localization is concerned: in one dimension, coherent back-scattering always strongly localizes the quantum states of the electrons; in three dimensions, electronic states are spatially extended (the electron ‘sea’ of familiar metals). 2D systems are more complicated. At a finite temperature, their electronic length scale is generally not set by the sample size, but by the so-called inelastic scattering length, determined by energy exchange between the electrons and their environment. This is temperature dependent, diverging to infinity at zero temperature — which would imply a conducting state. But weak localization theory⁴ predicts a logarithmic divergence of the resistance as the temperature is lowered; and experiments on silicon transistors and other 2D systems have confirmed this effect⁵.

However, weak localization theory does not explicitly consider the effect of the electrostatic Coulomb interaction between electrons. A generalization of the theory has been attempted by many theorists, particularly by Finkel’shtein⁶. Surprisingly, he found that interaction effects should make the ground state a perfect conductor, not an insulator. But the scaling procedure leads to a growth in the effective strength of the interactions, so this perturbative treatment breaks down close to zero temperature. This theoretical problem was not resolved, and so the experiments that pointed to the apparent success of weak localization theory seemed to settle the issue.

But in the past few years, doubt has re-emerged. Kravchenko and co-workers^{1,2} experimented on high-mobility samples (samples with very weak randomness), and varied the electron density. At low density, where interaction effects are large in comparison to the kinetic energy, the resistance diverges, as expected. But above a critical density, the resistance decreases with

Another surprise from two dimensions

T. Maurice Rice

Once again, after the discoveries of the quantum Hall effect and high-temperature superconductivity, it is a two-dimensional system that provides a real surprise in condensed-matter physics. For some years now there has been a general belief, bolstered by experiment, that the always-present randomness in a two-dimensional metal should cause its resistance to rise as the temperature is lowered, and inevitably lead to an insulating ground state. This view is now being challenged by a growing body of experimental evidence which points towards the existence of two phases at zero temperature, an insulator and a metal^{1,2}.

Furthermore, there are signs that the metallic ground state may be a perfect conductor, and an experiment reported last month³ shows that this effect disappears if the spins of the electrons are polarized — a clear sign that this is no ordinary metal.

The technology behind the semiconductor revolution allows one to fabricate an ideal system for fundamental investigations — a fluid of electrons, confined to move in a plane, whose density can be controlled externally and which is subject to only a weak random external potential. Theory began to catch up some twenty years ago, when four theorists (Abrahams, Anderson, Licciardello

decreasing temperature, tending towards a finite or even zero value at zero temperature.

How does this fit in with theory? A few months ago, Dobrosavljević and co-workers⁷ showed that scaling theory allows only two possibilities for the zero-temperature behaviour: either infinite resistance or infinite conductance (the latter caused by a many-body effect arising from the Coulomb interaction between the electrons). If this is true, the transition observed in experiments must be from perfectly conducting to perfectly insulating.

This picture is reinforced by the new study on the effect of a parallel magnetic field. Simonian and collaborators³ found that an external field of more than one tesla, applied in the plane of the sample, suppressed the metallic state completely. Such coplanar fields only polarize the spins of the electrons, which at higher temperatures and in ordinary metals does not lead to any drastic change in conductance. So the spin state is central to the high conductance of the metallic state — a characteristic of superconductors. In fact, Finkel'shtein⁵ and others⁸ showed that the interaction effects can be reliably treated for spin-polarized electrons, and predicted that only an insulating ground state is possible in this case.

Why might the metallic state be a perfect conductor at zero temperature? Is it just a standard superconductor, perhaps with an unusual internal symmetry⁹? Or is it a more unusual superconductor driven by a combined disorder and interaction effect¹⁰? Another exciting possibility is a form of superconductor with broken time-reversal symmetry, which would optimize the inter-electron correlation and so perhaps be stabilized by Coulomb interactions¹¹.

Further experiments will be a big help here. For example, by going to temperatures much lower than the present limit of 0.1 K, we can test the hypothesis of perfect conduc-

tance, and look for a finite-temperature superconducting phase transition. For now, this apparently simple system of a two-dimensional electron fluid moving in a weak random potential still poses some basic challenges to our understanding. □

T. Maurice Rice is in the Institute of Theoretical Physics, Eidgenössische Technische Hochschule, CH-8093 Zürich, Switzerland.

1. Kravchenko, S. V., Kravchenko, G. V., Furneaux, J. E., Pudalov, V. M. & D'Iorio, M. *Phys. Rev. B* **50**, 8039–8042 (1994).
2. Kravchenko, S. V. *et al. Phys. Rev. B* **51**, 7038–7045 (1995).
3. Simonian, D., Kravchenko, S. V., Sarachik, M. P. & Pudalov, V. M. *Phys. Rev. Lett.* **79**, 2304–2307 (1997).

4. Abrahams, E., Anderson, P. W., Licciardello, D. C. & Ramakrishnan, T. V. *Phys. Rev. Lett.* **42**, 673–676 (1979).
5. Bishop, D., Tsui, D. C. & Dynes, R. C. *Phys. Rev. Lett.* **44**, 1153–1156 (1980).
6. Finkel'shtein, A. M. *Sov. Phys. JETP* **57**, 97–108 (1983); *Sov. Phys. JETP Lett.* **40**, 796–799 (1984); *Z. Phys. B* **56**, 189–196 (1984).
7. Dobrosavljević, V., Abrahams, E., Miranda, E. & Chakravarty, S. *Phys. Rev. Lett.* **79**, 455–459 (1997).
8. Castellani, C., Di Castro, C., Lee, P. A. & Ma, M. *Phys. Rev.* **B30**, 527–543 (1984).
9. Phillips, P. & Wang, Y. preprint cond-mat/9704200 on xxx.lanl.gov.
10. Belitz, D. & Kirkpatrick, T. R. preprint cond-mat/9705023 on xxx.lanl.gov.
11. Ren, Y. & Zhang, F. C. *Phys. Rev. B* **49**, 1532–15350 (1994).

Palaeontology

Deep roots for the Neanderthals

G. Philip Rightmire

Since their bones were first discovered in the last century, the inventory of Neanderthals from Europe has grown substantially. These people are known from caves and rock shelters of the Late Pleistocene (130,000 years ago), and their anatomy has been described in detail. Older fossils that seem to predict the specialized Neanderthal morphology are also on record, and this lineage can be traced back in time for almost 200,000 years. What happened earlier is less clear, because fewer bones have been available for study. But this situation has now changed — since the mid-1980s, palaeontologists investigating an ancient cave near Burgos in northern Spain have uncovered a remarkable cache of human fossils. This assemblage from the Sima de los Huesos ('Pit of Bones') includes several well-preserved crania, along with jaws, teeth and other body parts of at least 32 individuals. The site and its contents are described by Juan-Luis Arsuaga and his colleagues¹ in the *Journal of Human Evolution*, and the finds reveal much about the people who preceded the Neander-

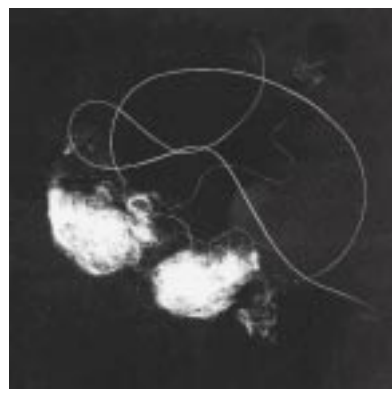
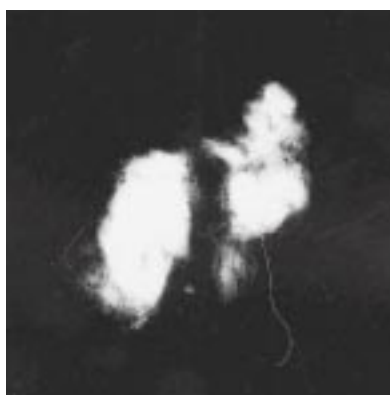
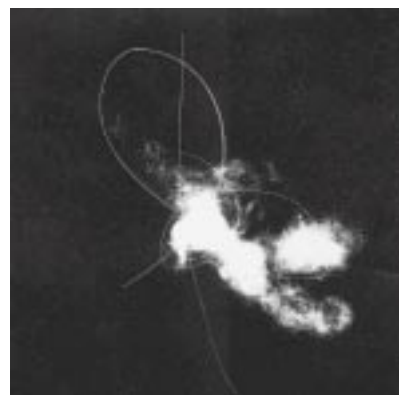
thals, and the course of human evolution in the Middle Pleistocene.

The Sima de los Huesos is part of a larger cave system in the Sierra de Atapuerca. Human remains predominate in the lower deposits of the pit itself, whereas bear bones have been excavated from overlying levels. Although there is some mixing of the bears with humans, animals other than carnivores are not well represented in the assemblage, and there are no artefacts. This suggests that the cave was not used as a living site. Some of the bears may have been trapped inside the cave later, and they probably disturbed or chewed on many of the bones². But how the hominid fossils came to be accumulated there is something of a mystery. Arsuaga *et al.*¹ favour an explanation involving the people themselves — they may have dumped whole bodies in the chamber as part of an early mortuary practice.

In any case, the bones are ancient³. Both human and bear fragments have been dated using electron-spin resonance and analysis of uranium isotopes. The results indicate a

Avoided object

Another matter Navel fluff: Sailor (while at sea), farmer, architect



This is a corrected version of a piece previously published.

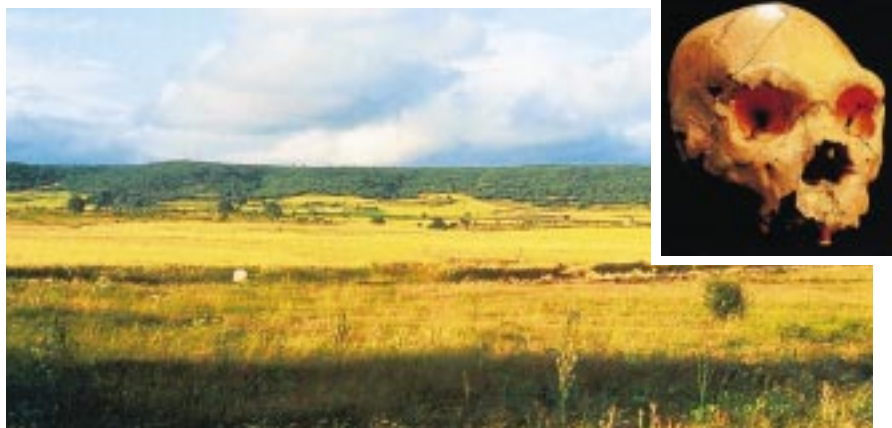


Figure 1 The Sima de los Huesos site in northern Spain, where human body parts dating back at least 200,000 years have been found. Among them is Cranium 5 (inset), one of the most complete fossil skulls of an archaic human ever found.

minimum age of about 200,000 years, although some of the remains could be in excess of 300,000 years old. As might be expected for specimens of such antiquity, the Sima crania show many primitive features that are found in other early Europeans, and in Middle Pleistocene skulls from Africa and Asia⁴. There are also specialized traits that suggest a link to the Neanderthals. The face is large and projecting in its middle parts and, in one well-preserved individual, the bone below the eye socket is slightly hollowed rather than flattened and continuous with the side wall of the nose. Thus, the cheek is not 'inflated' in the extreme manner of Neanderthals, although it may foreshadow this morphology. Also, brow ridges are very thick in the Atapuerca population, and the continuous form of these structures is reminiscent of Neanderthals. Additional resemblances are found in the anatomy of the temporal bone (which contributes to the side wall, ear region and base of the cranium), and in the occurrence of a shallow depression on the surface of the occiput (which makes up the rear of the braincase).

These clues are interpreted as evidence for evolutionary continuity in Europe. The Sima people seem to be related to Neanderthals. But in some respects they are intermediate between Neanderthals and still older fossils from Arago Cave in France, Bilzingsleben and Mauer in Germany, and other localities in Italy and Hungary. Although the Spanish researchers feel that there is little basis for splitting this long lineage into discrete segments, they elect to lump the Sima material — along with the more ancient remains — in the species *Homo heidelbergensis*. This taxon is said to be confined to Europe, where it is ancestral (only) to the Neanderthals. According to this 'accretion' hypothesis, Neanderthal distinctions accumulated gradually in populations that became isolated as a consequence of glacial conditions⁵. During the Middle Pleistocene (780,000–130,000 years ago), ice sheets to

the north and east, and the Mediterranean to the south, would have restricted gene exchange with people outside Europe. This would have led to full expression of the morphology that characterizes Late Pleistocene *H. neanderthalensis*.

Much of this hypothesis seems sound. Certainly, the link between the Atapuerca fossils and later populations is well established. Here the Sima de los Huesos opens a unique window on human evolution in the Middle Pleistocene, and there is now good reason to suggest deep roots for the Neanderthals. However, the extent to which older specimens from Arago, Bilzingsleben or Mauer show the special features of this lineage has still to be clarified, and some workers find it difficult to separate the earliest Europeans from contemporary people in Africa or even the Far East⁶. This raises the question of how *H. heidelbergensis* should be defined. It can be argued that the species was wide ranging and not restricted to Europe. Some populations were ancestral to Neanderthals but, perhaps in Africa, there was evolution in the direction of *H. sapiens*. In this version of events, *H. heidelbergensis* is the stem from which both Neanderthals and modern humans are derived. Choosing between these phylogenetic alternatives is difficult at present, and new evidence from both the Sima and other Atapuerca sites^{7,8} will have to be considered carefully as the issue is resolved. □

G. Philip Rightmire is in the Department of Anthropology, State University of New York at Binghamton, Binghamton, New York 13902-6000, USA.

e-mail: gpright@bingvmb.cc.binghamton.edu

1. Arsuaga, J.-L. et al. *J. Hum. Evol.* **33**, 109–127 (1997).
2. Andrews, P. & Fernandez Jalvo, Y. *J. Hum. Evol.* **33**, 191–217 (1997).
3. Bischoff, J. L. et al. *J. Hum. Evol.* **33**, 129–154 (1997).
4. Arsuaga, J. L., Martinez, I., Gracia, A. & Lorenzo, C. *J. Hum. Evol.* **33**, 219–281 (1997).
5. Hublin, J.-J. *Archaeology* **49**, 36–44 (1996).
6. Groves, C. P. & Lahr, M. M. *Persp. Hum. Biol.* **4**, 1–11 (1994).
7. Carbonell, E. et al. *Science* **269**, 826–830 (1995).
8. Bermudez de Castro, J. M. et al. *Science* **276**, 1392–1395 (1997).

Daedalus

Galactic whirlpools

Many astronomers resent the interstellar medium. It blots out their view of many interesting objects, such as the alleged black hole at the centre of our Galaxy. Daedalus sees it more positively.

He points out that all known galaxies are rotating. They spin, not like a wheel but like a bath-tub vortex. The matter near the centre travels faster than that further away — as Kepler would dictate. Yet their collective motion is still too wheel-like for comfort. To fit it to gravitational theory, astronomers must assume that each galaxy contains not only visible stars, but a vast outer ballast of invisible material. This is the mysterious 'dark matter'.

But Daedalus recalls that the viscosity of a gas is independent of its pressure. The tenuous interstellar medium should thus be about as viscous as normal air. On this basis, the vortex-like appearance of spiral galaxies is no surprise. A galaxy is indeed a vast bath-tub vortex, and its outer regions are rotated, not by the gravity of dark matter, but by viscous drag from the fast-spinning centre. He even muses that the vortex itself may be maintained by the inflow of stars and gas as they go down the awful 'plug-hole' of the central galactic black hole.

What happens to the energy viscously absorbed by the stirred interstellar medium? Daedalus calculates its Reynolds' number as well inside the turbulent region. Its energy will descend a whole hierarchy of swirls and eddies on its way to heat. This explains the puzzling heterogeneity of the medium, and the amazing temperatures it can attain — some parts reach 10^6 K, far hotter than any stellar surface. Furthermore, as in a bath-tub vortex, the turbulence must release a rude noise in some extreme low-frequency band. This will propagate as infrasound through the interstellar medium.

Infrasound originating nearby should be easiest to observe. Tenuous as it is, the interstellar medium must have a slight optical density. As its waves pass between us and the stellar background, the stars should dim and brighten slightly with its changing density. Their frequency will be very low — a cycle a year or even longer. Daedalus feels they could be best detected by continuous photon counting, for years on end, of specific star clusters (to discount individual stellar fluctuations). Data will trickle in very slowly; but who can say what advances might come from the new discipline of acoustic astronomy?

David Jones

A fascination with the principles behind structures led the Russian artist Naum Gabo to produce sculpture that looks as if it owes as much to engineering as to nature.

Intuition of the resonances between natural and human structures is as old as the oldest surviving writings on architecture — even older if we take into account Aristotle's fascination with nature's 'constructivists', such as the cell-building bee and geometrically accomplished spider.

A key motif for the early theorists and apologists of abstraction in the first half of this century was an appeal to the mathematical structures underlying natural appearance, as they were being revealed by science — or, in the Russian sculptor Naum Gabo's case, to the principles at work in the structures. Such an appeal was particularly favoured in the circles of British abstractionists, who continued to stand in the native tradition of nature-based art. It was this tradition with which Gabo interacted when he joined the colony of artists in St Ives, England, in 1939.

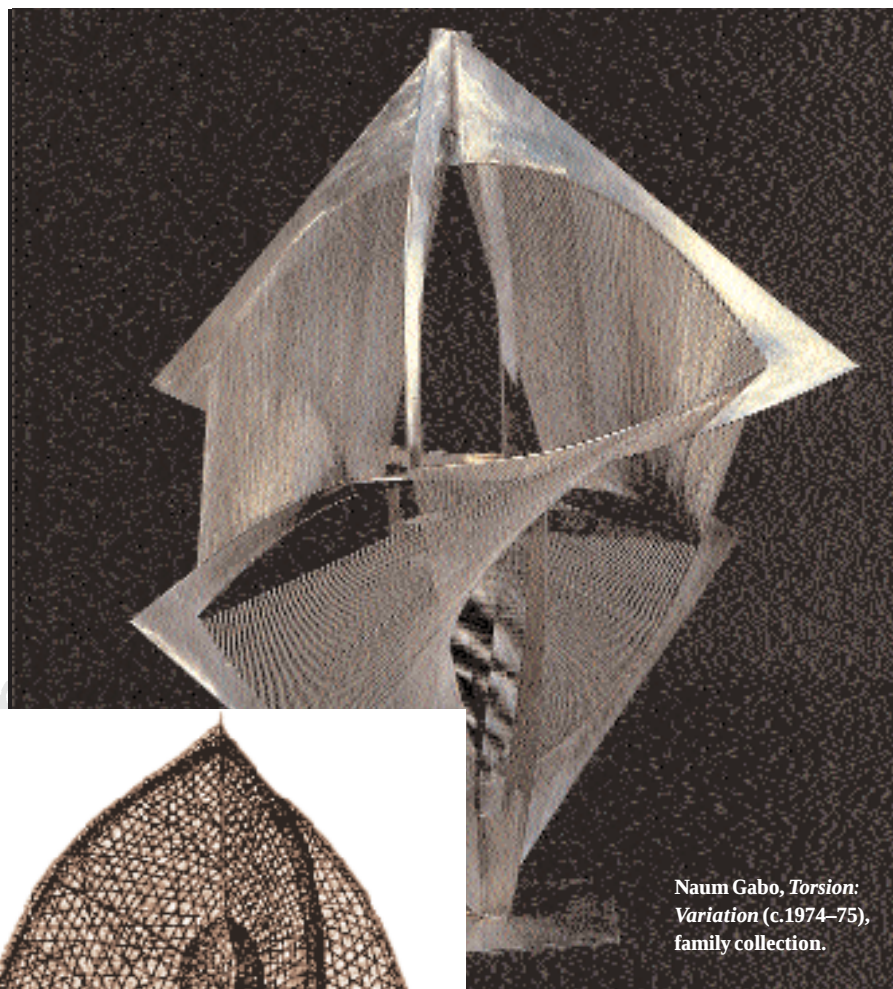
Educated as an engineer, Gabo had already formulated a credo based on nature's structural artifice. In *The Realist Manifesto* which he composed in 1920 with his brother Antoine Pevsner, he explained the principle on which they constructed their sculptural

"With a plumb line in the hand, with eyes as precise as a ruler, with a spirit as taut as a compass, we build them the same way as the universe builds its own creations, as the engineer his bridges, as the mathematician the formulae of his orbits."

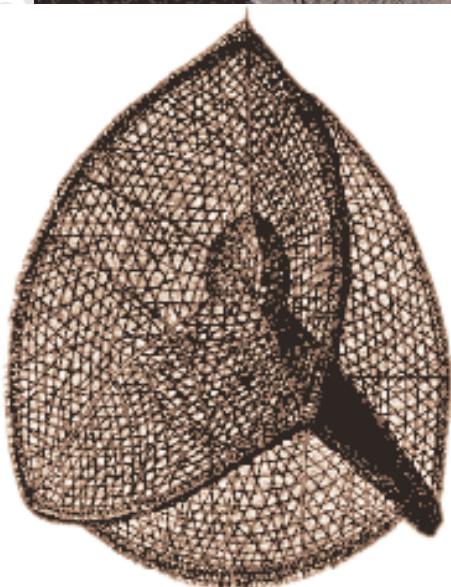
Given such a predilection, he was well placed to benefit from D'Arcy Wentworth Thompson's masterpiece of scientific writing, *On Growth and Form* (1917), to which he was introduced by the critic, Herbert Read. In Thompson he could read beguiling analyses of the structural and dynamic geometry of biological morphology in context of physical forces, both internal and external to the organism.

He could read of the microscopic miracles of natural design, such as the Nasselarian skeletons illustrated so beautifully in Ernst Haeckel's publications, not least his two-volume *Kunstformen der Natur* (1899–1904). Thompson drew Gabo's attention to the suggestive analogies between the skeleton of *Callimitra agnesae* and the curved films formed when Joseph Plateau dipped a tetrahedral cage into a soap solution.

A sculpture such as *Torsion: Variation*, tensely constructed from a stainless steel



Naum Gabo, *Torsion: Variation* (c.1974–75), family collection.

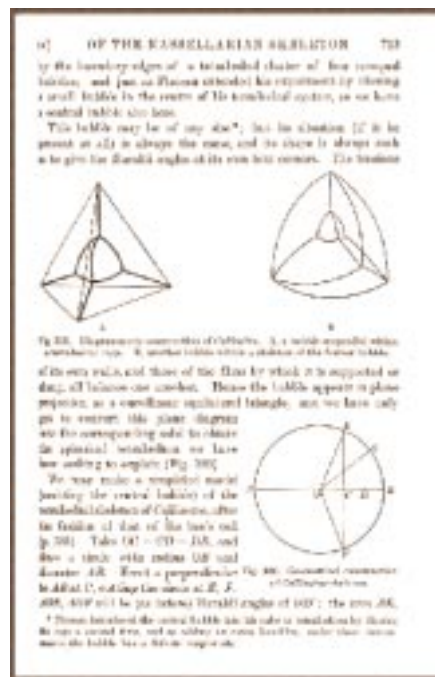


Skeleton of *Callimitra agnesae* (above) and its construction in terms of bubbles suspended in wire cages (right), from *On Growth and Form*.

exoskeleton with internal filaments of spring wire, can be seen as directly influenced by Thompson.

And in a broader sense it manifests the artist and scientist sharing an intuitive and intellectual communion with the statics of natural form at a profound level. $\square$

Martin Kemp is in the Department of the History of Art, University of Oxford, 35 Beaumont Street, Oxford OX1 2PG, UK.



Burrowing saves Lake Erie clams

Freshwater unionid clams in North America have been virtually eliminated from waters that are colonized by zebra mussels. Near total mortality has been reported in western Lake Erie^{1–4}, but we have now discovered a large population of native clams in a Lake Erie wetland that shows little sign of infestation. Field observations and laboratory experiments show that warm summer water temperatures and soft, silt-clay sediments trigger burrowing by clams. This discourages infestation and physically removes any attached zebra mussels.

Vegetation at Metzger Marsh, a 360-hectare wetland along the shore of western Lake Erie, is being restored to counter the effects of high lake levels and loss of a barrier beach by building of a dike and water-control structure to mimic the protective function of the beach. Surveys of biota before dike construction identified a large population of zebra mussels and five live unionids at the site. However, partial drainage exposed a larger clam population than expected.

We collected and relocated 6,000 unionids of 21 species, a similar diversity to that found in the open waters of Lake Erie before the zebra mussel invasion⁴, although most species have never been reported as occurring in wetlands. The size (10–240 mm) and age ranges (1–40 yr) of the population indicate that this wetland habitat has successfully supported clams for many years. In addition, every adult female examined was in reproductive condition. We found most unionids in a soft, silt-clay substrate (average organic content 10% and grain size mostly less than 500 µm), under about 1 m of water. Live unionids were not found in areas with coarse, sand-gravel substrates.

Zebra mussels initially invaded this area around 1990, but less than 1% of the unionids found were encrusted with zebra mussels or showed any sign of previous infestation. This indicates that there might be some type of behavioural mechanism by which clams either remained separate from drifting zebra mussel larvae or were able to remove attached zebra mussels. We found that unionids at Metzger Marsh burrowed 2–40 cm into the sediment for at least part of the day, and hypothesized that the soft, silt-clay sediments encouraged burrowing and so protected the clams from infestation.

To test this idea we placed twenty-five thick-shelled *Amblema plicata* and thin-shelled *Leptodea fragilis* of each of two size classes (shell length less than 60 mm and

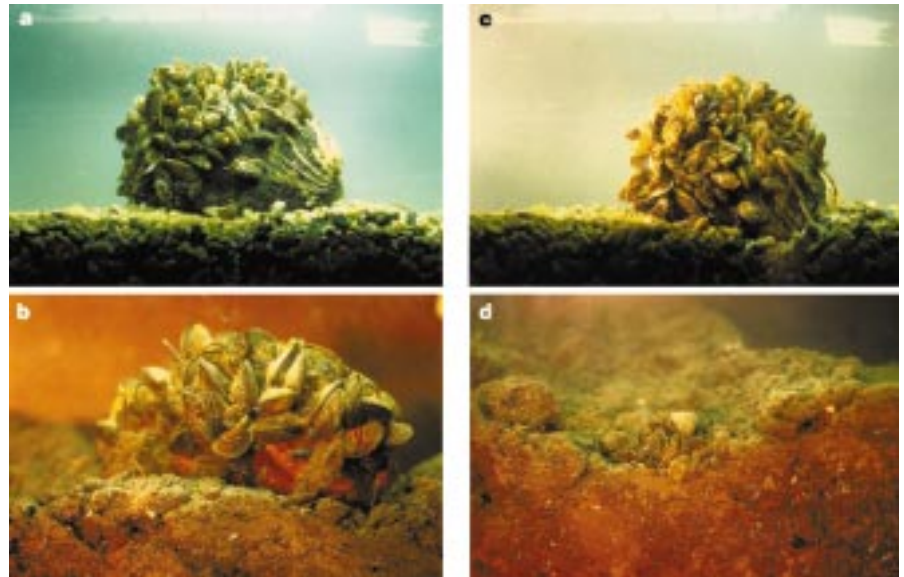


Figure 1 Burrowing of *Amblema* clams encrusted with zebra mussels. **a,b**, Encrusted clams were placed on Metzger Marsh sediments. **c**, After 24 h on coarse sand-gravel, clams burrowed only to the sediment contact with the zebra mussels. **d**, On the soft silt-clay, clams burrowed completely, also burying the zebra mussels.

more than 120 mm) in aquaria that contained either soft sediment or coarse sand-gravel from Metzger Marsh. At ambient temperatures (22 °C), there was no difference in the burrowing behaviour of clams in the two substrates. Small clams of both species burrowed completely within 4 h of being placed in the aquaria and large clams burrowed less than 10 mm. At 27 °C, consistent with normal summer temperatures in Metzger Marsh, large clams of both species burrowed as rapidly and as deeply as small clams in both substrates.

We tested large *Amblema* specimens (>120 mm) that were encrusted with zebra mussels, collected from Lake Michigan in Green Bay, Wisconsin, under the same conditions in aquaria at 27 °C. The number of zebra mussels on each of 15 animals ranged from 20 to 150. The encrusted clams were unable to burrow successfully in coarse sand-gravel (Fig. 1a, c). The clams burrowed until the first layer of zebra mussels came in contact with the sand but could burrow no further.

The same clams burrowed completely when placed in the soft sediments, carrying zebra mussels under the sediment with them (Fig. 1b, d). Most of the buried zebra mussels died after 24 hours, probably as a result of their inability to tolerate low levels of oxygen⁵, as the mortality rate dropped dramatically when the sediments were aerated. The movement of the native clams in and out of the soft sediment also dislodged small clusters of zebra mussels

attached to shells.

Native clams seem to have been protected from zebra mussel infestation at Metzger Marsh by the interaction between warm temperatures and soft sediments. Warm water encourages burrowing, but soft sediments are required to allow encrusted clams to burrow. In support of this conclusion, we have since found live unionids or fresh shells at three other Lake Erie wetlands.

These results provide promise that at least some brood stock might be available to recolonize Lake Erie if zebra mussel populations ultimately decline. Wetlands may provide a place for intensive management of native clam stocks, ensuring survival of these animals in the Great Lakes and other regions invaded by zebra mussels⁶.

S. Jerrine Nichols
Douglas A. Wilcox

Great Lakes Science Center,
US Geological Survey,
1451 Green Road,
Ann Arbor,
Michigan 48105, USA
e-mail: s.jerrine_nichols@usgs.gov

1. Strayer, D. L. & Smith, L. C. *Freshwater Biol.* **36**, 771–779 (1996).
2. Hunter, R. D. & Bailey, J. F. *Nautilus* **106**, 60–67 (1992).
3. Schloesser, D. W., Nalepa, T. F. & Mackie, G. L. *Am. Zool.* **36**, 300–310 (1996).
4. Schloesser, D. W. & Nalepa, T. F. *Can. J. Fish. Aquat. Sci.* **51**, 2234–2242 (1994).
5. Matthews, M. A. & McMahon, R. F. *Proc. 4th Int. Zebra Mussel Conf.* 231–249 (1994).
6. <http://www.glsc.nbs.gov>

Mixed blessings for middle-aged mothers

Perls *et al.*¹ in Scientific Correspondence present intriguing data suggesting that bearing children in middle age may be a harbinger for extreme longevity. We have reanalysed another study, which does not lead to the same conclusion.

Perls *et al.* report a significant fourfold excess of births in middle age ($N=19$ women age 40 or older) among 78 female centenarians when compared with a group of 54 women in the same birth cohort (born in Massachusetts in 1896) who survived only to the age of 73. We re-examined this question using data derived from a large case-control of postmenopausal breast cancer in Massachusetts, New Hampshire and Wisconsin². In that study, women provided detailed information on family history, including maternal and sibling dates of birth, survival status and dates of death, and data on socioeconomic status, race, religion and maternal smoking.

We identified the 3,578 mothers of control women (subjects without breast cancer) who survived at least to the age of 70 (70% were deceased at the time of the daughter's interview). All were born between 1867 and 1923 (median, 1900), and 552 (15.4%) had given birth after their fourth decade. In an unadjusted model, the mortality rate from all causes was slightly increased (by 11%) in association with late childbirth (rate ratio, 1.11; 95% confidence interval, 1.01–1.23; $P=0.04$). However, after adjusting for smoking and socioeconomic status, this excess disappeared (rate ratio, 1.01; 95% confidence interval, 0.90–1.13; $P=0.83$). Only 27 women lived to age 100 or beyond. Among the 1,167 women born in 1896 (the birth year of women in ref. 1) or earlier, the proportion bearing children after age 40 was higher in the 352 women who died between the ages of 70 and 79 (31%) than in the 322 women who survived to age 90 (24%). Thus, in this population of elderly parous women, all of whom had given birth to at least one child, we find no evidence that childbearing in 'middle age' is a reliable marker of longevity.

**Kathleen M. Egan, Edward Giovannucci
Linda Titus-Ernstoff, Polly Newcomb
Meir Stampfer**

Department of Epidemiology,
Harvard School of Public Health,
677 Huntington Avenue,
Boston, Massachusetts 02115, USA
e-mail: kathleen.egan@channing.harvard.edu

1. Perls, T. T., Alpert, L. & Fretts, R. C. *Nature* **389**, 133 (1997).
2. Newcomb, P. A. *et al.* *Am. J. Epidemiol.* **142**, 788–795 (1995).

There are several explanations for the evolution of human longevity and senescence^{1,2}, a recent example being the suggestion by

Perls *et al.*³ of a relationship between longevity in women and the age at which they bore children. Perls *et al.* showed that women surviving to at least age 100 were four times more likely to have had children while in their 40s than women who survived to age 73. But it is not possible to argue from these data that "the driving selective force of human lifespan is maximizing the period of time during which women can bear children" without a comparison of the total number of offspring produced (lifetime reproduction, LR) of the two groups of women.

Natural selection favours genes that increase LR^{2,4}, and consequently analyses must be based on LR if they are to conclude that the length of the reproductive period in women is a driving force in the evolution of human lifespan. The interpretation of Perls *et al.* would be supported if women in the centenarian group had a larger LR than those who died at 73. Alternatively, if women in the second group had equal or larger LR than those in the other, the shorter lifespan is probably an expression of costs associated with child care at an earlier age.

Philipp Heeb, Martin W. G. Brinkhof

Zoology Department, University of Bern,
CH-3032 Hinterkappelen, Switzerland
e-mail: philipp.heeb@esh.unibe.ch

1. Neese, R. M. & Williams, G. C. *Evolution and Healing* (Weinfielden & Nicholson, London, 1995).
2. Stearns, S. C. *The Evolution of Life Histories* (Oxford Univ. Press, 1992).
3. Perls, T. T., Alpert, L. & Fretts, R. C. *Nature* **389**, 133 (1997).
4. Daan, S. & Tinbergen, J. M. in *Behavioural Ecology* 4th edn (eds Krebs, J. R. & Davies, N. B.) 311–333 (Blackwell, Oxford, 1997).

In relation to the question of whether 'middle-aged' mothers live longer, raised recently in Scientific Correspondence¹, there are of course many factors that influence longevity. Menopause is not only a reflection of ovarian ageing: it could primarily be a result of hypothalamic disturbances, including neurotoxic effects of endogenous oestrogens². Further, later menopause is usually considered to be a risk factor for breast and endometrial cancer^{3,4}.

Another factor affecting lifespan is smoking, which may strongly influence the timing of the onset of menopause, age at the time of last birth and longevity⁵. The connection between giving birth after age 40 and lifespan may be modulated by the number of times a woman has previously given birth⁶. And the lifespan of female progeny of women giving birth for the last time after the age of 40 may be shorter than that of the progeny of younger mothers⁷. Additional investigation is needed before a definite conclusion can be drawn about the relationship between longevity and the age of parturition.

Lev M. Berstein

Laboratory of Endocrinology,
N. N. Petrov Research Institute of Oncology,
St Petersburg 189646, Russia
e-mail: levmb@endocrin.spb.ru

1. Perls, T. T., Alpert, L. & Fretts, R. C. *Nature* **389**, 133 (1997).
2. Brawer, J. R. *et al.* *Biol. Reprod.* **49**, 647–652 (1993).
3. MacMahon, B. *et al.* *J. Natl Cancer Inst.* **50**, 21–42 (1973).
4. Brinton, L. A. *et al.* *Am. J. Obst. Gynecol.* **167**, 1317–1325 (1992).
5. Berstein, L. M. *Oncoendocrinology of Smoking* (Nauka, St Petersburg, 1995).
6. Hytten, F. E. & Leitch, I. *The Physiology of Human Pregnancy* (Blackwell Scientific, Oxford, 1971).
7. Gavrilov, L. A. *et al.* *Proc. Acad. Sci. (Russia)* **354**, 569–572 (1997).

Tree-ring dating the 1700 Cascadia earthquake

Geological evidence shows that an earthquake attended by a tsunami, or a series of such earthquakes, ruptured at least 900 km of the Cascadia subduction zone along the west coast of North America between the years 1700 and 1720¹. Satake *et al.*² found evidence for one large tsunami of previously unknown origin in Japanese archives of that era, and used its time and size to propose that a Cascadia earthquake of magnitude 9 occurred on 26 January 1700. Tree-ring records from a central part of the Cascadia subduction zone now provide support for this proposal.

Much of the Cascadia coast between southern British Columbia and northern California was lowered abruptly as a result of an earthquake (or earthquakes) about 300 years ago^{1,3}. In Washington state, coastal forests were submerged by more than a metre of tidewater⁴, damaging some trees and killing thousands of others. Several dozen Sitka spruce (*Picea sitchensis*) that survived this tidal submergence remain around the estuaries of southern Washington. Annual growth rings in half of the trees show changes in width and anatomy consistent with disturbance (tilting, increased flooding, or both) in the first few years after 1699⁵.

We instead used annual rings to date western red cedar (*Thuja plicata*) killed by

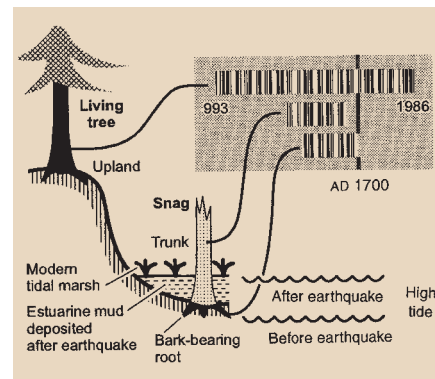


Figure 1 Approach for tree-ring dating of earthquake-induced submergence. Upland trees contain ring-width patterns that correlate with those of dead trees (snags) killed by tidal submergence. Final rings in snags have been lost on weather-beaten trunks but are preserved in bark-covered roots. Bar codes exaggerate ring-width correlations.

Table 1 Dating of red cedar snags that retain intact roots

Estuary	Snag	Trunk (or root, PX-J6 only)			Root		
		Sample length (yr)	Outer preserved ring (year AD)	Correlation with Long Island r , log P	Sample length (yr)	Outer ring against bark (year AD)	Correlation with trunk r , log P
Copalis River	CP-791	283	1680	0.28, -3.42	ND	1708	T
	CP-GF2	255	1675	0.21, -0.76*	331	1699 L	0.57, -21.39
Grays Harbor	JN-560	295	1674	0.26, -2.71	ND	1699 L	T
	JN-561	263	1678	0.44, -10.56	248	1699 L	0.35, -5.97
Willapa Bay	PX-782	239	1681	0.52, -15.16	355	1699 L	0.42, -9.63; T
	PX-783	340	1675	0.30, -5.59	218	1699 L	0.63, -20.71; T
	PX-J6	202	1699	0.23, -0.54*	ND	1699	ND
Columbia River	GR-777	378	1671	0.27, -4.55	90	1699 L	T

r , correlation coefficient; P , probability of error in correlation; T, rings independently correlated by tracing from trunk to root. L, 1699 ring contains latewood. ND, not determined. Assumptions: trunks, 200-yr minimum overlap and a latest-possible date of 1720 based on radiocarbon; roots, outer rings postdate those of trunks but are no later than 1720. Data from trees CP-GF2 and PX-J6 were from G. Jacoby. Samples were collected with L. Amidon, P. Atwater, K. Bevis, J. Boughner, S. Brown, M. Cisternas, S. Cooley, L. Davis, S. Fritts, N. Jacobson, J. Morales, P. Raskind, R. Lewis, J. C. Moya, K. Sayce, J. Shulene, F. Taylor, Y. Wang, K. Wegmann and A. Zachery. *Correlations with securely dated snags at the same estuary (assuming a 150-yr minimum overlap): CP-GF2, $r = 0.28$, log $P = -3.18$; PX-J6, $r = 0.29$, log $P = -2.16$.

the submergence (Fig. 1 and supplementary information). We first identified consistent ring-width patterns in 19 upland trees on Long Island, Willapa Bay (studied with C. Woodhouse). These red cedar are rooted about 50 m above sea level, too high to have suffered from the submergence. Their composite ring-width pattern extends from AD 993 to 1986. Next we looked for segments of the Long Island pattern in the decayed trunks of 75 earthquake-killed trees (snags) in tidal wetlands of the Copalis River, Grays Harbor, Willapa Bay, and the Columbia River, along 90 km of the subduction zone. We identified such segments in 15 of the snags. The probability of error in these matches is less than 0.001 for 13 of the snags and less than 0.01 for the other two (examples in Table 1). We dated six additional snags by correlating them with these firmly dated trees.

Trunk rings from the last years before death have been lost because of exposure, but the dated trunk rings show that many of the red cedars died after the 1670s and 1680s. At least one snag from each of the four estuaries retains a trunk ring from 1680 or later, and a Columbia River snag extends to 1691.

To date tree death more exactly we used bark-bearing roots of the snags, in collaboration with G. Jacoby. Of 12 snag roots sampled, eight gave reliable estimates of the final year of tree growth (Table 1). Concentric rings in roots of these eight snags can be counted as far as the bark along radii that had been the most rapidly growing before death. In seven of the eight snags, including one or more at each of the four estuaries, the final ring of the sampled root formed in 1699. The other snag gave an outer-ring date of 1708, but the exceptional height of its dated root may have delayed death from tidal submergence.

Six of the snags narrow the time of tree death to the months between the end of the 1699 growing season (August) and the start of the 1700 growing season (May). In these snags, the final ring, which dates to 1699,

contains dark, late-summer cells where the ring is more than a few cells wide (L, Table 1). In the seventh snag, which has a final ring from 1699, the ring is too thin to show whether such latewood is present or not.

By converging so narrowly on January 1700, these tree-ring dates securely link an earthquake inferred from geology at Cascadia with a tsunami known from documents in Japan. They thereby give the 1700 Cascadia earthquake a place in written history even though it predates the region's earliest documents⁸ by almost a century.

The dates provide a simple test of earthquake size. Suppose that the dates excluded January 1700. In that case, the tsunami in Japan could not represent a 650-km-long rupture at Cascadia — the distance between the region of the dated snags and the far (California) end of the subduction zone. Because a 650-km-long rupture of the Cascadia would be too small for magnitude 9 (ref. 7), dates excluding January 1700 would thereby strengthen geophysical arguments⁸ against magnitude 9 earthquakes at Cascadia. By converging on January 1700, the dates mean that Canada and the northwestern United States are plausibly subject to earthquakes of magnitude 9.

David K. Yamaguchi

Department of Environmental Health,
University of Washington, Box 354695,
Seattle, Washington 98195, USA
e-mail: yamaguch@u.washington.edu

Brian F. Atwater

US Geological Survey,
Department of Geological Sciences,
University of Washington, Box 351310,
Seattle, Washington 98195, USA

Daniel E. Bunker

Tree-Ring Laboratory,
Lamont-Doherty Earth Observatory,
Palisades, New York 10964, USA

Boyd E. Benson

GeoEngineers Inc., 8410 154th Avenue N.E.,
Redmond, Washington 98052, USA

Marion S. Reid

The Nature Conservancy, 2060 Broadway,
Suite 230, Boulder, Colorado 80302, USA

- Nelson, A. R. *et al.* *Nature* **378**, 371–374 (1995).
- Satake, K., Shimazaki, K., Tsuji, Y. & Ueda, K. *Nature* **379**, 246–249 (1996).
- Atwater, B. F. *et al.* *Earthquake Spectra* **11**, 1–18 (1995).
- Hemphill-Haley, E. *Geol. Soc. Am. Bull.* **107**, 367–378 (1995).
- Jacoby, G. C., Bunker, D. E. & Benson, B. E. *Geology* **25**, 999–1002 (1997).
- Heaton, T. H. & Hartzell, S. H. *Science* **236**, 162–168 (1987).
- Hyndman, R. D. & Wang, K. J. *Geophys. Res.* **100**, 22133 (1995).
- McCaffrey, R. & Goldfinger, C. *Science* **267**, 856–859 (1995).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from Mary Sheehan at the London editorial office of Nature.

Thermal contraction

The Scientific Correspondence by Roy and Agrawal¹ commenting on a News and Views article by Robert Cahn² states that their work had not been acknowledged by Cahn or by us. Cahn's article emphasized negative thermal expansion, but the three reviews mentioned by Roy and Agrawal are about low thermal expansion. All of our papers (refs 3–5 in ref. 2) prominently mention the work of Roy and Agrawal on the family of compounds based on $\text{NaZr}_2\text{P}_3\text{O}_{12}$ (NZP). Further, the thermal expansion properties in the zirconium tungstate family^{3–5} highlighted by Cahn² differ dramatically from other materials, including those of the NZP family. Roy and Agrawal stated that the unusual thermal expansion in the NZP family was discovered by them, but this was first seen by Boilot *et al.*⁶, as they acknowledge in their earlier papers⁷.

Possibly the most misleading part of Roy and Agrawal's contribution concerns a particular member of the NZP family, $\text{NaTi}_2\text{P}_3\text{O}_{12}$. They give its thermal expansion as -5.5×10^{-6} per °C, from measurements of the expansion of a ceramic body. In fact, the intrinsic linear thermal expansion of this material, on the basis of unit-cell changes with temperature, is $+4.3 \times 10^{-6}$ per °C (ref. 8).

Thermal expansion in the NZP family is usually highly anisotropic, as indeed it is for $\text{NaZr}_2\text{P}_3\text{O}_{12}$ itself, the thermal expansion of individual crystals being strongly positive in one dimension and negative in two dimensions. The effects of anisotropic thermal expansion of crystallites on a ceramic body composed of such crystallites have long been known⁹. The inherent microcracking generally decreases mechanical strength, causes hysteresis in thermal expansion properties, changes the thermal expansion properties with thermal cycling, and causes a decrease in thermal expansion observed in ceramic bodies because the cracks tend to become larger with decreasing temperature. Ceramic bodies of $\text{NaTi}_2\text{P}_3\text{O}_{12}$ normally show a thermal expansion less negative than the -5.5×10^{-6} per °C value quoted by Roy and Agrawal. However, in using a material to lower the thermal expansion of a composite, it is the intrinsic thermal expansion

and other properties, such as bulk modulus, which are important. The highest intrinsic negative thermal expansion reported in the NZP family is about -1×10^{-6} per °C. The thermal expansion for ZrW_2O_8 , on the other hand, is -8.7×10^{-6} per °C.

Roy and Agrawal also suggested that the thermal expansion mechanism in ZrW_2O_8 is the same as in the NZP family. Actually, the thermal expansion of cordierite ($\text{Mg}_2\text{Al}_2\text{Si}_2\text{O}_{10}$), β -eucryptite (LiAlSiO_4), and NZP is dominated by the thermal expansion of Mg–O, Li–O or Na–O bonds⁵. Such materials never show a strong negative thermal expansion of the unit-cell volume.

The mechanism for thermal expansion in ZrW_2O_8 is very different from that of the NZP family². In the ZrW_2O_8 family, transverse thermal motion of oxygen atoms causes negative thermal expansion, the same mechanism that causes negative thermal expansion in various forms of silica over limited temperature ranges that do not include room temperature. The surprising property of the ZrW_2O_8 family is that negative thermal expansion occurs over such a wide temperature range.

Roy and Agrawal also comment on the use of anisotropic materials in applications. It is true that those at Corning have made good use of anisotropic materials, such as cordierite ($\text{Mg}_2\text{Al}_2\text{Si}_2\text{O}_{10}$), for low thermal expansion applications, but it has been generally appreciated that isotropic materials would be favoured for most applications⁹. The anticipated link between anisotropy and microcracking has been verified experimentally in the NZP family¹⁰, and those working on the family considered it a major achievement when they found very low anisotropy in the $(\text{Ca},\text{Sr})\text{Zr}_4\text{P}_6\text{O}_{24}$ system.

The thermal expansion properties in the ZrW_2O_8 family are unlike those of any other known materials. Their intrinsic negative linear thermal expansion over a broad temperature range (including room temperature) is much higher than reported for any other material. This, and the strictly isotropic behaviour, open up many potential applications that could not be contemplated with other materials, such as those of the NZP family.

Arthur W. Sleight

Department of Chemistry, Oregon State University,
Corvallis, Oregon 97331, USA
e-mail: sleight@chem.orst.edu

Gene flow from transgenic crops

Gene flow from crops to related wild species must be considered when assessing the potential environmental impact of cultivating genetically modified plants¹. Evidence of pollen dispersal within species has been found for several crops but little information is available on spontaneous gene flow from crops to related species with simultaneous flowering periods^{2,3}. To study the genetic mechanisms involved, we have developed an intergeneric model of gene flow from transgenic oilseed rape (*Brassica napus* L.; genotype, AACC; diploid chromosome number, $2n=38$) containing one copy of the *bar* gene, which confers resistance to the herbicide Basta (glufosinate ammonium), to wild radish (*Raphanus raphanistrum* L.; genotype, RrRr; $2n=18$), a widely distributed weed.

We obtained oilseed rape $\times$ wild radish F_1 interspecific hybrids⁴ and studied the four successive generations under field conditions. The hybrids were surrounded by wild radish plants.

All oilseed rape varieties used as mother plants produced seeds. In the first generation, female fertility of the F_1 interspecific hybrids was poor, with only 28.6% producing seeds. The number of seeds was ten times higher at the second generation, and the increased female fertility was confirmed with the third-generation plants. The seed set (ranging from 0 to 5,673 seeds per plant) was, in some plants, close to the observed fertility of wild radish plants (500 to 10,000 seeds per plant; Table 1). In all generations, the percentage of seed germination never limited the production of successive progeny, which ranged from 74.2 to 88.9%.

In agreement with previous data⁴, 99.5% of first-generation interspecific hybrids had the expected genomic structure, ACRr, $2n=28$, that is, half of the genome of each parent. Of the second-generation hybrids, 48% corresponded to the structure, ACRrRr, $2n=37$. All others had either more chromosomes (33.4% of the total, mainly showing an amphidiploid structure, AACCRrRr, $2n=56$) or less chromosomes

(18.6% of the total, mainly showing the same chromosome number as the mother plant, ACRr, $2n=28$). Female fertility was highly dependent on the chromosome number of the mother plant. Mother plants with $2n=56$, $2n=37$ and $2n=28$ produced from 0.9, 8.4 and 59.7 seeds per plant, respectively. At the third generation, chromosome number generally decreased, with 83.6% of the plants having less than 28 chromosomes. The plants with the lowest chromosome numbers were the most fertile. We confirmed this tendency towards chromosome decrease in the fourth-generation plants, with 89.5% of the plants having less than 27 chromosomes. We observed a chromosome number close to that of wild radish ($2n=18$) in 25.4% of the plants (Table 1).

Basta resistance in the F_1 interspecific hybrids displayed mendelian segregation, that is, a 1:1 ratio of resistant and susceptible plants, as the oilseed rape mother plants were heterozygous for the *bar* transgene. However, because of unreduced gametes, the *bar* gene transmission was high in first-generation interspecific hybrids. The *bar* gene transmission was dependent on the chromosome number of the mother plant and decreased in successive generations (Table 1).

It seems that intergeneric gene flow might mainly occur by transgene introgression within the genome of the weeds, but slowly and at a low probability under natural optimal conditions because four generations were needed to provide herbicide-resistant plants with a chromosome number and morphology close to that of the weed. It is likely that under normal agricultural conditions this event is rare when the wild radish is the female parent⁵.

**Anne-Marie Chèvre, Frédérique Eber
Alain Baranger, Michel Renard**

Station d'Amélioration des Plantes,
INRA BP29, F 35653 Le Rheu Cedex, France

1. Dale, P. J. & Irwin, J. A. *Euphytica* **85**, 425–431 (1995).
2. Chèvre, A. M. et al. *Acta Hort.* **407**, 169–179 (1996).
3. Mikkelsen, T. R., Andersen, B. & Jørgensen, R. B. *Nature* **380**, 31 (1996).
4. Baranger, A., Chèvre, A. M., Eber, F. & Renard, M. *Theor. Appl. Genet.* **91**, 956–963 (1995).
5. Darmency, H., Fleury, A. & Lefol, E. *Brighton Crop Protect. Conf. Weeds* 433–438 (1995).
6. Eber, F., Letanneur, J. C. & Chèvre, A. M. *Cruciferae Newslett.* **19**, 17–18 (1997).

Table 1 Gene flow analysis from transgenic oilseed rape to wild radish

	Female fertility				Chromosome number				Basta resistance	
	No. plants	Seeds per 100 flowers	Seeds per plant	Plants with ≥ 1 seed	No. plants	Range $2n$	Mode $2n$	Plants in mode (%)	No. plants	Resistant plants (%)
Male-sterile oilseed rape	589	43.2	29.2	100	589	(38)	[38]	100	589	100
Generation 1	1,301	0.2	11	28.6	2,057	(28–56)	[28]	99.5	2,619	51.8
2	251	1.5	11.9	42.2	686	(25–84)	[37]	48.0	674	81.9
3	443	7.9	229.3	82.5	1,083	(18–60)	[24–27]	45.3	1,028	57.2
4					787	(18–59)	[21–23]	45.3	3,349	23.5

Chromosome number was established either by mitotic chromosome counting or by flow cytometry^{4,6}. The presence of the transgene was checked by spraying herbicide solution on 3–4-leaf-stage seedlings⁴.

1. Roy, R. & Agrawal, D. *Nature* **388**, 433 (1997).
2. Cahn, R. *Nature* **386**, 22–23 (1997).
3. Korhonen, V. et al. *Chem. Mater.* **7**, 412–417 (1995).
4. Evans, J. S. O., Mary, T. A., Vogt, T., Subramanian, M. A. & Sleight, A. W. *Chem. Mater.* **8**, 2809–2823 (1996).
5. Sleight, A. W. *Endeavour* **19**, 64–68 (1995).
6. Boilot, J. P., Salanie, J. P., Desplanches, G. & Le Potier, D. *Mater. Res. Bull.* **14**, 1469–1477 (1979).
7. Agrawal, D. K. & Roy, R. *J. Mater. Sci.* **20**, 4617–4623 (1985).
8. Oota, T. & Yamai, I. *J. Am. Ceram. Soc.* **69**, 1–6 (1986).
9. Buessem, W. R. in *Mechanical Properties of Engineering Ceramics* (eds Kreigel, W. W. & Palmour, H. III) 127–148 (Interscience, New York, 1961).
10. Srikanth, V. *J. Am. Ceram. Soc.* **74**, 365–368 (1991).

Reasoning in the social sciences

The World Made New: Frederick Soddy, Science, Politics and Environment

by Linda Merricks

Oxford University Press: 1997. Pp. 223. £60, \$105

A. G. Maddock

A Nobel laureate in chemistry and a man with a very strong social conscience, Frederick Soddy (1877–1956) is a rewarding subject for a biographer. For more than a decade at the beginning of the century he was a brilliant and productive natural scientist. Then, from shortly after the First World War to the end of his life, he forsook these studies for a rather unsuccessful foray into political and economic sciences.

His researches in his earlier period, at McGill, Glasgow and Aberdeen, complemented those of his friend Ernest Rutherford. Together they established the nature of radioactivity, the atom and isotopes. At times Soddy's achievements have been underestimated, although they never were by Rutherford. His standing was probably not enhanced by an over-adulatory account by Muriel Howarth published in 1958. Nevertheless, everyone who attended secondary school since the end of the Second World War is acquainted with his results, although they may not associate them with Soddy's name.

Soddy was always interested in the use society made of discoveries in science. Unlike Rutherford, he believed that a way would be found to tap the enormous energy associated with nuclear forces. Indeed, H. G. Wells' extraordinarily prescient novel, *The World Set Free*, stems from reading Soddy's book *The Interpretation of the Radium* (1912).

An inspiring lecturer, his written work was always presented in a popular form. His acknowledged dislike of a mathematical approach to natural science may sometimes have been responsible for his work failing to be properly appreciated.

In 1919 he was appointed as one of the Dr Lee professors of chemistry at Oxford. (The other was F. A. Lindemann.) It undoubtedly true that Soddy was a difficult and disputatious man. But he was certainly treated rather badly by Oxford. For his first five years in this post he had neither laboratory accommodation nor a budget for laboratory expenses. A technicality even excluded him from membership of the faculty board.

Linda Merricks has carried out a great deal of research into Soddy's arguments and difficulties during this period and throughout his life. As in many biographies, undue emphasis is at times placed on what may have been trivial incidents or explosive initial reactions to a disappointment. But certainly these difficul-



Frederick Soddy: gifted natural scientist with a mild persecution complex.

ties interrupted his research and may have led him to devote his time to the politics of the social application of science.

The author traces Soddy's involvement with various political societies and pressure groups. Soddy's ability as a lecturer led to his considerable influence on his contemporaries, but by the time J. D. Bernal was writing his classic work on the history of science he made no mention of Soddy.

Soddy's studies led quite naturally to an interest in economics, particularly the basis of money, and by 1930 he had lost practically all interest in the exact sciences. Although Merricks gives a detailed account of the people and bodies Soddy was involved with, she does not explain his ideas: indeed, she remarks that it was difficult to discover exactly what they were. So the book fails to address the question of what Soddy's activities in politics and economics achieved.

Soddy had a very happy married life, and the death of his wife in 1936 had a profound effect on him. It was followed the next year by the death of his friend Rutherford. He resigned from his Oxford chair and his life became dark and joyless until the end of the Second World War.

For several years science, politics and economics ceased to interest him and he took refuge in mathematical studies. Much of this work was confined to notebooks, but he did publish two or three papers on the fringe of mathematics. The lack of a mathematical approach in his scientific papers was a matter of choice rather than due to lack of ability.

Merricks suggests that Soddy's empirical approach to research became unsuitable after the development of quantum mechanics, but Hahn, Hevesy and Paneth and others were successful with a similar approach.

Soddy found relaxation in climbing and hill walking, developing a great liking for the Lake District. He was always fond of travel and in 1937 he revisited Ceylon, a country that had much attracted him during a voyage to Australia in 1904. The trip momentarily rekindled his interest in science and he took the opportunity to collect some monazite sand samples from Travancore.

The Second World War was inimical to all Soddy's beliefs and he played no part in it. But the dramatic ending with atomic bombs stimulated him to renewed activity, both scientific and political. His belief in nuclear power and the need for control of scientific discoveries had been vindicated. He attended scientific meetings again, and in 1949 produced the first European account of atomic power written by a scientist not involved in the wartime projects. Unfortunately, the book was poorly advertised and indeed difficult to buy, so Soddy received little credit.

The lack of appreciation of his work led to a mild persecution complex, and it is hard to escape the conclusion that such personal frustrations contributed to a large degree to a life that was, for all its early achievements, curiously unfulfilled. □

A. G. Maddock is at 4 Grange Court, Pinehurst, Grange Road, Cambridge CB3 9BD, UK.



Underwater sex change

Wrasses, members of the Labridae family, are abundant inhabitants of the coral reefs around the Arabian Peninsula. The many species range from a few centimetres in length to more than 2 metres in the case of the largest, the humphead wrasse, *Cheilinus undulatus* (above). Some members of the family can change sex, their

coloration altering gradually during the transformation. From *Desert Sea: Fauna of the Saudi-Arabian Red Sea Coast* by Hagen Schmid (distributed in the United Kingdom by Gazelle Book Services, £28.99). The book contains many fine photographs but only a brief text.

The data day

The Illustrated Almanac of Science, Technology, and Invention: Facts, Figures, and the Fanciful
by Raymond L. Francis
Plenum: 1997. Pp. 376. £17.55, \$28.95
John Heilbron

Did you know that in 1994 an okapi at Copenhagen Zoo died of an overdose of Wagner? So says Raymond L. Francis, who reports that the delicate animal preferred death to listening to local singers rehearsing Tannhäuser. It thus showed greater discrimination than Francis, who, as the report suggests, stuffs his almanac with bits as foreign to it as opera to an okapi.

Every day in the year has its own page, with a dozen or so entries from various years and, usually, an apt figure or cartoon. The death of the okapi appears on 5 August and, again, on the sixth — an achievement made possible, it must be assumed, by the author's unwillingness to read through his finished work, for which he can scarcely be blamed. Other double entries occur for J. J. Silvester and W. F. Libby, in each case including a full repetition of information and anecdotes. No doubt there are many more.

The want of method indicated by this repetition and irrelevance can be further illustrated. There are four entries on the McDonald's hamburger chain, two on singing telegrams, and one each on the abdication of the last Queen of Hawaii, the establishment of the Swiss Guard at the Vatican, the first scalping of an Indian by a white man, the beginning of the Pony Express and the foundation of the city of Rome. Also mentioned are the opening of the first hotel in the United States with indoor toilets, the thousandth known suicide from the Golden Gate Bridge, the highest high jump of a dog, and the sale of the skull of Emanuel Swedenborg.

The entries dealing with science and technology come largely from the most obvious sources: notices of scientists (mainly Nobel prizewinners) and engineers; accounts of discoveries, especially ones with several datable episodes (Galileo's career, the travels and writings of Darwin and Wallace, the discovery and follow-up of X-rays and Becquerel rays, the announcements of nuclear fission); patents; voyages, notably Columbus's and NASA's; and modern almanacs, conspicuously those for 1993 and 1994. In ransacking this material, Francis has given particular attention to conservation and the envi-

ronment, which generated the most interesting and useful entries in his book.

One cannot cover so large and amorphous a field without making mistakes. A few may be signalled as indications of the level at which the entries are written and their occasional shortfalls. Heisenberg did not receive the Nobel prize for the uncertainty principle nor Blackett for "his (atom) smashing research"; Planck did not say in 1900 that radiant energy comes in quanta; Marat, murdered 13 July 1793, did not arrange for the execution of Lavoisier, guillotined 8 May 1794; the 'feud' between Galileo and the Roman Catholic Church was not about sunspots; and John Dee did not make the English translation of Euclid to which he contributed his famous preface.

Francis's jumble is worth dipping into for a moment's amusement. It may be useful to editors needing fillers for columns on "this day in history" and to seekers after odd coincidences, such as the almost simultaneous (modulo 365) cessation of the longest sneezing fit in history and the shooting of *The Sneeze*, the first film copyrighted in the United States. True anniversarilists will gain little from the almanac, however, because it has no listing by year (though a good index of names). Only by slogging through it can information relevant to centennials be gathered. What has been culled in this laborious manner for anniversaries due for celebration next year will appear in the first number of *Nature* for 1998. □

John Heilbron is senior research fellow at Worcester College, Oxford, and at the Museum of History of Science, Broad Street, Oxford OX1 3AZ, UK.

Duke of hazards

Perils of a Restless Planet:
Scientific Perspectives on
Natural Disasters
by Ernest Zebrowski
Cambridge University Press: 1997. Pp. 306.
£16.95, \$24.95
Bill McGuire

Whether it is because it is the International Decade for Natural Disaster Reduction, as designated by the United Nations, or a result of the approaching millennium, natural hazards are once again in vogue. As a consequence, the 1990s have spawned a plethora of texts on the subject, many to underpin the growing numbers of environmentally based degree courses across the world. Unfortunately, most of these texts are much less exciting than their subject matter, with stock hazards addressed in a dry, descriptive and unimaginative way, and often recycling fact and fiction from similar texts of earlier generations. It is therefore refreshing to come across a new

approach to natural disasters, especially one written in an easy yet informative style.

Ernest Zebrowski's brief — or at least one of them — is to highlight the science behind natural hazards and their disastrous consequences, rather than simply describing their effects. So the chapter on, say, the impact of earthquakes introduces the reader to concepts of both material strength and forces. Discussion of tsunamis is underpinned by a consideration of wave characteristics and their propagation, and that of hurricanes and windstorms by variations in the forces transmitted to buildings owing to their shape.

The opening chapter provides the reader with a real feel for the awesome power of nature. In a fascinating account of the great earthquake that devastated Lisbon in 1755, we learn that the city was not only battered by a magnitude 8.75 quake — one of the largest known — but that it later succumbed to a devastating tsunami and then to a conflagration. Then comes a description of the demise of the Minoan civilization as the island of Santorini blew itself apart during the great eruption of 1640 BC.

After this exciting opening, however, the author changes tack dramatically, dedicating the next chapter to explaining — in some detail — the history and philosophy of science from Aristotle onwards. Although a good read, this chapter sits slightly uneasily in the text as a whole, and explicit links with natural disasters are tenuous to say the least.

Returning to the principal theme of the book, the third chapter examines the problems of living in buildings in earthquake zones and illustrates how their safety is a function of the nature and strength of the materials used in their construction, the way the buildings are put together, and the forces applied during a quake. Strangely, this does not lead into a look at some of the main earthquake events and their effects, but into a second, idiosyncratic chapter that

focuses primarily on the characteristics of population growth. In essence, the message is clear: the more people there are, the more who are likely to die in natural disasters.

The author then goes on to provide informative and scientifically based accounts of destructive waves and important earthquakes and their effects. Surprisingly, in the treatment of destructive waves, he pays little attention to the scale of tsunami generated by large collapses from island volcanoes, probably the largest and potentially most destructive waves possible, barring those caused by impact events in the oceans.

Which brings me conveniently to the next chapter. For some reason, the author deals with volcanoes and asteroid impacts together, leading to a dilution of coverage of these two most devastating of natural hazards. Despite the destructive power of floods, earthquakes and windstorms, only volcanic eruptions and impact events have the power to cause catastrophe on a global scale. The absence of any reference to major volcanic eruptions is therefore surprising. After all, the eruption of Toba some 74,000 years ago lowered temperatures in some parts of the world by more than 10 °C, and has been proposed as the final straw that launched an already cooling planet into full ice age.

Following an account of windstorms and their effects, the closing chapter addresses the concept of the planet as a chaotic system and the inevitable scenario of a flitting butterfly ultimately generating a hurricane half a world away. Intuitively, such a situation still fails to convince me. Surely, the buffering effect of countless other atmospheric disturbances would drown out such an infinitesimally small event well before it led to the mass destruction of Miami?

With its strong personal perspective, this book stands out from the recent clutch of disaster tomes. Although idiosyncratic from time to time, and perhaps a little illog-

ical in organization and construction, it provides an original insight into natural disasters and their causes. It will be a refreshing read for hazard scientists and students, as well as the world at large. □

Bill McGuire is at the Greig Fester Centre for Hazard Research, Department of Geological Sciences, University College London, Gower Street, London WC1E 6BT, UK.

Methods and madness

Mental Ills and Bodily Cures: Psychiatric Treatment in the First Half of the Twentieth Century
by Joel Braslow
University of California Press: 1997. Pp. 240. \$40, £30

John Burnham

"Mother has received a letter about this new brain operation. She has asked me to write to you as you prowle [sic] know how old folks feel about any new operation they can't seem to believe will do any good." In *Mental Ills and Bodily Cures*, Joel Braslow, who holds a doctorate in history and is a practising clinical psychiatrist, tells the story of how one somatic therapy succeeded another in psychiatry during the heyday of the insane asylum in the United States — and why. Although he draws specifically on material from two hospitals in California, the treatments and issues he discusses were international.

On the face of it, this book appears to be a modest historical monograph on the application of various therapies to certain patients in a relatively local setting. But the issues it explores address fundamental questions — even about the standards of what is scientific and how science should be applied, especially in medicine. For some critics of science, the development of psychiatric ideas and practices has provided a vulnerable target for their attempts to discredit the application of any naturalistic approach to human problems. Generations have grown up thinking that Ken Kesey's *One Flew Over the Cuckoo's Nest* — and in particular the film version — provides an accurate picture of psychiatric treatment in the first half of this century; and tendentious histories of psychiatric treatments provide examples that show why science and medicine are bad.

Braslow's account will therefore reverberate far beyond the bounds of medicine and history. Moreover, it is not only remarkably understandable, but also contains many case histories with often dramatic dialogue to grip the attention. And if ever an account deserved the currently clichéd accolade of 'nuanced', this one does: no individual instance, used to make a point, is separated



Goodbye waves: the destructive power of water as demonstrated in Johnstown, Pennsylvania, in 1889.



Insane treatment: Jack Nicholson in *One Flew over the Cuckoos Nest*.

from the scientific literature or statistical and epidemiological context. Those still touched by postmodernism will enjoy the irony of the story the author tells and even the three small — and moving — autobiographical snippets from his own life.

All intellectuals, whether trendy or traditional, will appreciate Braslow's consideration of the anti-psychiatry movement, particularly the branch opened up by the philosopher Michel Foucault. In response to others who explicitly try to use history for social purposes in bottom-up rather than top-down accounts, Braslow works hard to elicit the voice of the patient as well as the doctor. He has incorporated some of the viewpoints of these anti-psychiatry scholars. He addresses their questions — and then he trumps them.

Historians seeking to establish a narrative of modern medicine have moved back and

forth between science and therapy. Braslow focuses on modes of treatment, but he also recaps the standard scientific history of each innovation. Only after he has shown the original therapeutic rationale does he describe what happened as physicians and patients negotiated the use of each treatment.

By the end of the nineteenth century, physical restraint of patients was generally agreed to be undesirable — but nothing was as undesirable as uncontrolled behaviour, which often caused injury and was deeply distressing to the patient. ("She was given to violent attacks.... Restraint and room seclusion were necessary. Last June, when she became dangerous to keep in a dormitory unit, she was sent to the chronic disturbed unit.")

The first new somatic modality was hydrotherapy: wet packs and the continuous

bath, both of which often had calming effects. The second was sterilization, which in practice lost the eugenic purpose on which so many accounts have focused, and which, both doctors and patients believed, had desirable therapeutic effects on individuals. The next intervention was malaria fever treatment for tertiary syphilis of the brain (the effectiveness of which won a Nobel prize for Julius Wagner-Jauregg). Then in the 1930s came insulin and electroshock therapies. Finally, lobotomy was introduced, followed by the chemically invasive neuroleptics.

As one therapy followed another (and could disappear even more suddenly than it arrived), tendencies emerged in the succession of procedures. One line of progression was the shift of attention from the sick person to the disease — noted long since in general medical practice. Another continuity was the change in the way in which doctors and patients negotiated treatment, as the rationales for each type of therapy became increasingly 'scientific'.

A third continuity, and the one with which Braslow is justifiably most fascinated, is the way in which "the history of twentieth-century somatic therapies is a series of increasingly invasive attempts to control the psychiatric patient through his or her body". At first, in hydrotherapy, the attack on the body was superficial. Gradually, however, the invasion targeted the brain directly, until, with surgery and chemicals, it was increasingly directed towards specific areas of the brain and ultimately specific brain cells.

But each time psychiatrists provided a biological basis for their therapies, they tended to lose their patients — to surgery or to internal medicine (as in lobotomy and tertiary syphilis). Braslow shows that biological psychiatrists are today following a dangerous path, and he perhaps more than implicitly suggests that they should concentrate on the relationship between the doctor and the patient — exactly the direction in which historical reconstructions started moving a generation ago.

Braslow points out that there are no villains in the book, and indeed one of his great accomplishments is to vindicate the motives of physicians. There is irony and tragedy and much complexity here: it is a book about an imperfect world. Never mind that the author slights the overwhelming presence of the senile, interprets to the point of mind reading in places, and gratuitously plays with the Foucaultian idea of 'discipline'. His narrative decisively conveys the reality of psychiatric practice and illness from 1900 to 1954 more accurately than previous accounts. The moral to his story will weigh heavily on any thoughtful reader. □

John Burnham is in the Departments of History and Psychiatry, Ohio State University, 230 West

New in paperback

Without Miracles: Universal Selection Theory and the Second Darwinian Revolution

by Gary Cziko

MIT Press, \$17.50, £14.95

Reviewed by David Hull in *Nature* 377, 494 (1995).

The Simple Science of Flight: From Insects to Jumbo Jets

by Henk Tennekes

MIT Press, \$12.50, £10.50

Reviewed by C. J. Pennycuik in *Nature* 381, 126 (1996).

Artificial Minds

by Stan Franklin

MIT Press, \$17.50, £14.95

Reviewed by Terrence J. Sejnowski in *Nature* 378, 101 (1995).

Queer Science: The Use and Abuse of Research into Homosexuality

by Simon LeVay

MIT Press, \$15, £12.50

Reviewed by D. F. Swaab in *Nature* 382, 682 (1996).

The Evolution of Communication

by Marc D. Hauser

MIT Press, \$35, £29.50

Reviewed by Derek Bickerton in *Nature* 382, 592 (1996).

Autumn Books

Next week's issue will contain *Nature's* Autumn Books supplement, with reviews on the topics of time, eggs and sperm, French natural history, anti-reductionism, meta-analysis and much more.

Contribution of Southern Ocean surface-water stratification to low atmospheric CO₂ concentrations during the last glacial period

Roger François^{*}, Mark A. Altabet[†], Ein-Fen Yu[‡], Daniel M. Sigman[§], Michael P. Bacon^{*}, Martine Frank^{||}, Gerhard Bohrmann[¶], Gilles Bareille[#] & Laurent D. Labeyrie^{*}

^{*} Department of Marine Chemistry and Geochemistry, [§] Department of Geology and Geophysics, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

[†] Department of Chemistry and Biochemistry and Center for Marine Science and Technology, University of Massachusetts, North Dartmouth, Massachusetts 02747, USA

[‡] Department of Earth Sciences, National Taiwan Normal University, Taipei, 117 Taiwan

^{||} Department of Earth Sciences, University of Oxford, Oxford OX1 3PR, UK

[¶] GEOMAR Research Center for Marine Geosciences, 24148 Kiel, Germany

[#] Département de Géologie et Océanographie, Université de Bordeaux I, Talence, 33405 Cedex, France

^{*} Centre des Faibles Radioactivités, Laboratoire mixte CNRS-CEA, Gif-sur-Yvette, 91198 Cedex, France

The nitrogen-isotope record preserved in Southern Ocean sediments, along with several geochemical tracers for the settling fluxes of biogenic matter, reveals patterns of past nutrient supply to phytoplankton and surface-water stratification in this oceanic region. Areal averaging of these spatial patterns indicates that reduction of the CO₂ 'leak' from ocean to atmosphere by increased surface-water stratification south of the Polar Front made a greater contribution to the lowering of atmospheric CO₂ concentration during the Last Glacial Maximum than did the increased export of organic carbon from surface to deep waters occurring further north.

The unstratified, nutrient-rich surface waters in the modern high-latitude ocean provide the main conduit for transferring deep-water CO₂ back into the atmosphere. This CO₂ 'leak' to the atmosphere is particularly effective in the modern Southern Ocean, because of intensive vertical mixing and low nutrient utilization. Building on this observation, a series of papers^{1–6} attributed the glacial lowering of atmospheric CO₂ either to enhanced biological removal of the nutrients and CO₂ in high-latitude surface waters resulting in higher sinking fluxes of organic matter (that is, higher export production), or to a lower supply rate of nutrients and CO₂ from intermediate waters produced by lower vertical mixing. But the extensive deep- or intermediate-water anoxia predicted by these models, and the lack of palaeo-oceanographic evidence for increased export production in the glacial Southern Ocean, has led to the questioning of the validity of these models. Here we present new evidence which supports increased stratification south of the position of the modern Polar Front (MPF) during the Last Glacial Maximum as a mechanism that contributed to lower glacial atmospheric CO₂ and deep-water oxygen concentration.

The sedimentary record of several palaeoproductivity proxies was recently presented as support for a large increase in the export flux of organic carbon in the Atlantic sector of the Southern Ocean, north of the position of the MPF, which could have contributed to lowering glacial atmospheric CO₂ as a result of Fe fertilization⁷. Here we combine a similar suite of geochemical proxies for palaeoproductivity and its fate with bulk-sediment $\delta^{15}\text{N}$ values to better constrain past changes in the nutrient balance of surface waters and its influence on atmospheric CO₂. Bulk-sediment $\delta^{15}\text{N}$ provides a means of evaluating the fraction of surface nitrate utilized by phytoplankton^{8–11}. By combining this information with export flux of nitrogen estimated from palaeoproductivity proxies, we constrain the supply rate of nitrate to surface waters. Our results indicate that, despite higher export flux of organic carbon in the

Atlantic sector of the Southern Ocean north of the MPF during the last glacial period, the fraction of nitrate utilized by phytoplankton in this region did not increase, implying a sustained supply of nutrients (and thus CO₂) to surface waters. In contrast, south of the MPF, export production decreased in the Atlantic and western Indian sectors, or changed very little in the eastern Indian sector, while the fraction of nitrate utilized by phytoplankton increased significantly. Similar or lower export flux and increased biological utilization of nitrate imply a substantial decrease in the rate of supply of nutrients and CO₂ to surface waters throughout this region. These observations indicate that surface-water stratification, rather than increased export production, was the primary Southern Ocean process contributing to the lowering of glacial atmospheric CO₂.

Opal fluxes

Earlier estimates of past changes in Southern Ocean productivity were based on opal burial rates^{12–14} and were affected by post-depositional sediment redistribution by bottom currents. This problem has been alleviated by normalizing opal flux to excess ²³⁰Th activity in sediments, a method which largely corrects for sediment focusing or winnowing and provides estimates of preserved rain rate—pelagic settling fluxes minus dissolution rate on the sea floor^{15–17}. During the Holocene epoch, the preserved rain rates of opal are below the detection limits (BDL) north of the Subantarctic Front (Fig. 1a). The highest preserved opal rain rates (~2 g cm⁻² kyr⁻¹) occur in the Atlantic sector and on Kerguelen plateau in the Indian sector, slightly south of the MPF (~50° S), and correspond to the northernmost region with high silicate concentration in surface waters¹⁸. Although few results are yet available further south, preserved opal rain rates seem to decrease¹⁹. During the last glacial, ²³⁰Th-normalized opal fluxes show clearly the northwards migration of the frontal system by ~5° latitude

(Fig. 1b), as reported in earlier micropalaeontological^{20,21} and geochemical^{12–14,17} studies. In the Atlantic, the belt of high opal flux remained at $\sim 2 \text{ g cm}^{-2} \text{ kyr}^{-1}$ while migrating northwards, and resulted in a glacial increase of opal fluxes to the north of the MPF and a decrease to the south (Fig. 2a). In contrast, latitudinal changes in opal flux in the eastern Indian sector are much more subdued. These data suggest higher Holocene opal productivity in the Atlantic sector and on Kerguelen plateau, than in other regions south of the MPF, in a pattern comparable to satellite-derived regional distribution of annually averaged pigment concentration in surface waters²². They also suggest little change in the overall opal productivity of the Southern Ocean during the last glacial period, in agreement with earlier estimates^{12,17}.

Opal dissolution may have changed on the glacial/interglacial timescale, however, obscuring the interpretation of preserved opal flux in terms of opal productivity. Unlike ^{230}Th concentration in sediments and ^{230}Th -normalized fluxes, the ratio $\text{ex.}^{231}\text{Pa}_0/\text{ex.}^{230}\text{Th}_0$ is not affected by dissolution on the sea floor, and documents past changes in opal export fluxes rather than just preserved rain rates¹⁷ ($\text{ex.}^{231}\text{Pa}_0$ and $\text{ex.}^{230}\text{Th}_0$ are the activities of ^{231}Pa and ^{230}Th in excess of secular equilibrium with uranium present in the lattices of

minerals in sediment, decay-corrected to the time of deposition).

Because it has a longer residence time in sea water, ^{231}Pa is preferentially scavenged in high flux regions^{23,24} where settling materials and underlying sediments acquire an $\text{ex.}^{231}\text{Pa}_0/\text{ex.}^{230}\text{Th}_0$ ratio higher than their production-rate ratio from the decay of uranium dissolved in sea water (0.093). In contrast, sediments deposited in low-flux regions have a ^{231}Pa deficit (that is, $\text{ex.}^{231}\text{Pa}_0/\text{ex.}^{230}\text{Th}_0 < 0.093$). The relationship between $\text{ex.}^{231}\text{Pa}_0/\text{ex.}^{230}\text{Th}_0$ and particle flux is complex²⁵, however, and depends on the particle flux not only at the site of study, but also at neighbouring sites, on the rate of lateral transport of ^{231}Pa between sites, and on opal concentration in settling particles²⁶. In particular, when applying this approach to the Southern Ocean, advection of ^{231}Pa from the Atlantic Ocean with North Atlantic Deep Water must be taken into consideration, as $\sim 50\%$ of the ^{231}Pa produced in the Atlantic is exported into the Southern Ocean²⁷. As a result, accurate interpretation of $\text{ex.}^{231}\text{Pa}_0/\text{ex.}^{230}\text{Th}_0$ requires synoptic maps, at least on a basin-wide scale.

The distribution of $\text{ex.}^{231}\text{Pa}_0/\text{ex.}^{230}\text{Th}_0$ in Southern Ocean sediments follows closely the pattern of ^{230}Th -normalized opal fluxes. The ratio is below the production rate ratio (0.093) north of the

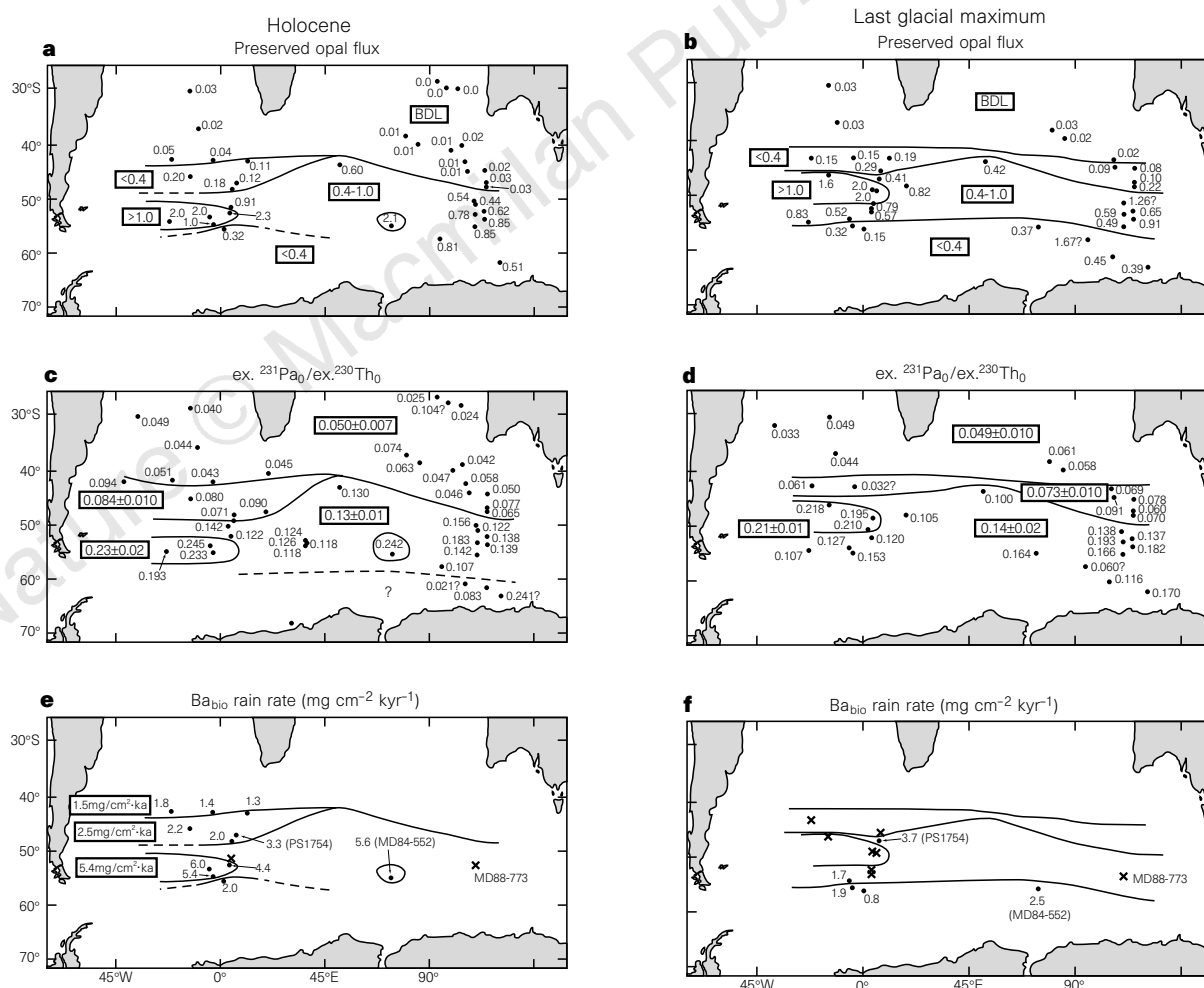


Figure 1 Comparison of geochemical data for Holocene and last-glacial times derived from deep-sea sediment cores. ^{230}Th -normalized opal fluxes in $\text{cm}^{-2} \text{ kyr}^{-1}$ (averaged values are given unless a single point analysis was performed; BDL indicates below detection limits for opal analysis) during the Holocene (**a**) and the last glacial (**b**). We also show $\text{ex.}^{231}\text{Pa}_0/\text{ex.}^{230}\text{Th}_0$ in Holocene (**c**) and last glacial (**d**) sediment, and dissolution-corrected ^{230}Th -normalized biogenic Ba fluxes during the Holocene (**e**) and the last glacial (**f**).

Export production of organic carbon (P_{exp}) was calculated according to ref. 31: $P_{\text{exp}} = 1.95 (\text{Ba}_{\text{bio}} \text{ rain rate})^{1.41}$. Crosses (in **e**, and **f**) indicate core sections with poor barite preservation based on high authigenic U activity ($> 0.3 \text{ d.p.m. g}^{-1}$) concurrent with high sediment accumulation rate ($> 10 \text{ cm kyr}^{-1}$). Boundaries in **e** and **f** reflect the geographical distribution of ^{230}Th -normalized opal fluxes (**a** and **b**). Values in boxes are the ranges (**a**, **b**) or the averages (**c**, **d**, **e**) within the areas delineated by the boundaries reported in each figure.

Subantarctic Front (Fig. 1c). It approaches 0.093 towards the Polar Front, coinciding with a gradual increase in opal preserved flux (Fig. 1a). During the Holocene, this transition zone is wider in the Atlantic sector. South of the Polar Front, the ratios are invariably higher than the production rate ratio. Clear maxima at 52°–55° S in the Atlantic and on Kerguelen plateau during the Holocene (Fig. 1c), and at 46°–51° S in the glacial Atlantic (Fig. 1d), document the northward shift of the frontal system. As for the ^{230}Th -normalized opal flux, relatively little glacial to Holocene change occurs in the eastern Indian sector (Fig. 2b). South of the high preserved-opal flux belt of the Atlantic, the ratio decreases (particularly during the glacial) but, at equal preserved opal fluxes, remains significantly higher than in the subantarctic region. This is due to the higher percentage of opal in settling particles, which decreases the partitioning coefficient between ^{231}Pa and ^{230}Th during scavenging²⁶.

Export of ^{231}Pa from the Atlantic to the Southern Ocean during the Last Glacial Maximum was similar to that of the Holocene²⁷. Consequently, last glacial to Holocene changes in the $\text{ex.}^{231}\text{Pa}_0/\text{ex.}^{230}\text{Th}_0$ of Southern Ocean sediment must primarily reflect relative changes in opal flux. The general similarity between $\text{ex.}^{231}\text{Pa}_0/\text{ex.}^{230}\text{Th}_0$ and ^{230}Th -normalized opal flux indicates that the latter is minimally biased by opal dissolution and primarily reflects relative changes in opal rain rate. Because of the variability in the opal to organic carbon ratio, however, changes in opal flux cannot be directly translated into changes in the export flux of organic carbon. Estimating the latter, which is necessary to understand the role of the Southern Ocean in controlling glacial atmospheric CO_2 , requires alternative approaches.

Sedimentary accumulation of authigenic uranium

In oxic pore water, U is present as a soluble carbonate complex. When conditions become sufficiently reducing to initiate iron or sulphate reduction, U is reduced to an insoluble form and precipitates, creating a porewater gradient, which causes U to move from the overlying sea water into the sediments²⁸. The U added to the sediment by this process (referred to as authigenic U) accumulates at a rate which is established by the porewater U concentration gradient, which depends on the level of oxygen in bottom water and the flux of metabolizable organic matter from both export production and sediment focusing.

Considering the very large authigenic U concentration in the glacial section of cores taken north of the MPF in the Atlantic sector,

it has been argued⁷ that the glacial increase in export flux of carbon in this region was much larger than suggested by the increase in opal flux, and similar to modern productive margins. Significant contribution to authigenic U addition from lower oxygen concentrations in bottom water was dismissed, on the basis of the absence of authigenic U in the glacial section of a core situated to the south (RC13-259; 53° 53' S, 3° 20' E). However, accumulation of authigenic U in sediments is the result of a complex, nonlinear interaction among the three primary controlling factors (that is, bottom-water oxygen concentrations, export flux of metabolizable organic matter, and sediment focusing), and the absence of authigenic U in this core cannot be taken as unambiguous evidence for a lack of decrease in bottom-water O_2 concentration. RC13-259 has a relatively low sedimentation rate, indicating little sediment focusing, and records low glacial productivity⁷. Deep-water O_2 concentration may still have decreased significantly, but the organic carbon input from export flux and focusing could have been below the threshold value required for creating the reducing conditions for authigenic U accumulation²⁹. In contrast, we have positive evidence for a lowering of O_2 concentration in bottom water from authigenic U accumulation in two cores from the Indian sector of the Southern Ocean (Fig. 3). In core MD84-552, authigenic U accumulates during isotopic stage 2, even though export-production proxies indicate lower export flux, while accumulation rates of $\text{ex.}^{230}\text{Th}_0$ indicate no sediment focusing during the glacial (a useful parameter to consider here is the focusing factor (Ψ), which is the ratio of decay-corrected ^{230}Th accumulation rate over production rate in the overlying water column¹⁵; $\Psi_{\text{Hol}} = 1.6$; $\Psi_{\text{Gl}} = 1.1$). In core MD88-773, a large increase in authigenic U is found for the last glacial, when export flux was similar to the Holocene and focusing was lower ($\Psi_{\text{Hol}} = 8.9$; $\Psi_{\text{Gl}} = 4.2$). Similar trends were found in Atlantic core PS1768-8¹⁹ (52° 36' S, 4° 28' E). These glacial maxima in authigenic U can be explained only by a decrease in bottom-water oxygen concentration, which would have contributed to the large authigenic U peaks recorded in the glacial Atlantic sector north of the MPF.

Export fluxes of organic carbon

A close link between annually averaged biogenic Ba fluxes ($\text{Ba}_{\text{bio}} = \Sigma\text{Ba} - 0.0075 \text{ Al}\%$) and export production was demonstrated by sediment-trap studies^{30,31}. ^{230}Th -normalized fluxes of Ba_{bio} thus have the potential to provide quantitative estimates of

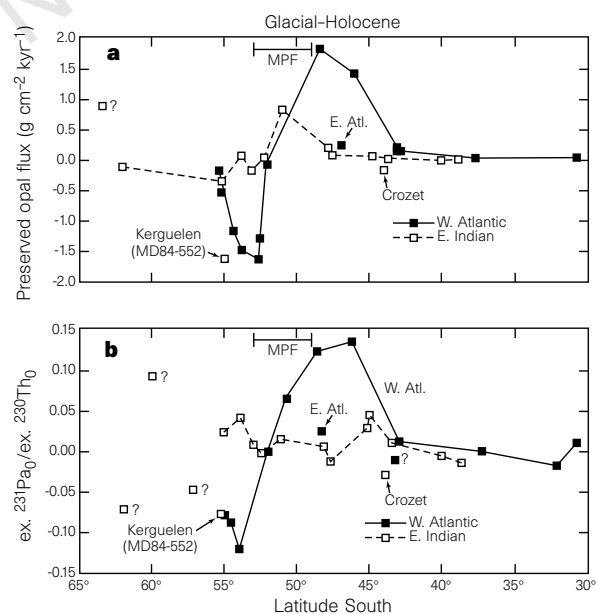


Figure 2 Difference between last-glacial and Holocene values of geochemical data. **a**, ^{230}Th -normalized opal fluxes; **b**, $\text{ex.}^{231}\text{Pa}_0/\text{ex.}^{230}\text{Th}_0$; values shown are last-glacial minus Holocene. Data from refs 7, 19, 25, 27, 49; G.B., unpublished data; and R.F., unpublished data. MPF indicates the range of latitudes at which the Polar Front is found in the modern Atlantic and Eastern Indian sector of the Southern Ocean.

export flux of organic carbon, if barite dissolution on the sea floor can be estimated. Dymond *et al.*³⁰ have found a first-order correlation between sediment accumulation rate and barite preservation. If we use their relationship to correct for barite dissolution, the dissolution-corrected ^{230}Th -normalized biogenic barium fluxes (or Ba_{bio} rain rates) in Holocene sediments show a clear maximum in the high-opal-flux belt of the Atlantic and Indian sectors, with values ranging between 4.4 and $6.0 \text{ mg cm}^{-2} \text{ kyr}^{-1}$ (Fig. 1e). Using the algorithm proposed by François *et al.*³¹, this would correspond to $15\text{--}24 \text{ g C m}^{-2} \text{ yr}^{-1}$ export production. North of the Polar Front, biogenic Ba rain rates gradually decrease to $1.4\text{--}1.8 \text{ mg cm}^{-2} \text{ kyr}^{-1}$ at $\sim 43^\circ \text{S}$ in the Atlantic, which correspond to $\sim 3\text{--}4 \text{ g C m}^{-2} \text{ yr}^{-1}$ export production. As expected, the northward decrease in biogenic Ba fluxes is not as pronounced as for opal fluxes, reflecting the transition from diatom to carbonate-producing plankton and the poor preservation of opal in sediments¹⁹.

Care must be taken, however, when interpreting biogenic Ba fluxes, as preservation may depend not only on sediment accumulation rates, but also on the redox conditions prevailing in the sediment. In high-accumulation-rate reducing sediments, rates of sulphate reduction could exceed rates of sulphate diffusion from bottom waters. Under these conditions, barite would start to dissolve, as evidenced by porewater profiles³². Ba fluxes obtained from high-sedimentation-rate cores with high authigenic U concentrations are thus unreliable, especially if Ba disagrees with other palaeoflux proxies. This is generally the case for the glacial section of cores collected in the Atlantic sector, north of the Polar Front (Fig. 1f), where glacial biogenic Ba rain rates fail to record the glacial increase in export production recorded by preserved opal flux and $\text{ex.}^{231}\text{Pa}_0/\text{ex.}^{230}\text{Th}_0$ (Fig. 2). Only one core from this area (PS1754¹⁹; Fig. 1f) has a sufficiently low glacial sedimentation rate (2.7 cm kyr^{-1}) and authigenic U content ($0.08 \text{ d.p.m. g}^{-1}$, disintegrations per minute per gram) to have a reliable biogenic Ba

record. Its biogenic Ba rain rate ($3.7 \text{ mg cm}^{-2} \text{ kyr}^{-1}$) indicates marginally higher export production during the glacial (12 versus $10 \text{ g C m}^{-2} \text{ yr}^{-1}$), to a level similar to that north of the modern high-opal belt (Fig. 1e).

Combining the Ba record from core PS1754 with evidence from opal preserved flux (Fig. 1a, b) and $\text{ex.}^{231}\text{Pa}_0/\text{ex.}^{230}\text{Th}_0$ (Fig. 1c, d) for changes in opal productivity, we conclude that glacial productivity north of the Polar Front increased to a level comparable to the region south of the Polar Front in the modern ocean, rather than to much higher levels similar to modern productive margins⁷. This interpretation is also consistent with evidence for lower O_2 concentration in glacial bottom water (Fig. 3) which implies that the high glacial authigenic U peaks north of the MPF⁷ do not reflect only higher export flux of organic carbon.

South of the MPF, four cores have comparatively low sedimentation rates and authigenic U. Their Ba records clearly show a glacial decrease in export production (Fig. 1f), in agreement with the other proxies. The Holocene biogenic Ba rain rate in the southernmost core (PS1772; $55^\circ 27' \text{S}$, $1^\circ 10' \text{E}$) is comparatively low ($2.0 \text{ mg cm}^{-2} \text{ kyr}^{-1}$) and similar to glacial values from the two Atlantic cores immediately to the north (1.7 and $1.9 \text{ mg cm}^{-2} \text{ kyr}^{-1}$), indicating low export production ($\sim 5 \text{ g C m}^{-2} \text{ yr}^{-1}$) and corroborating the preserved opal flux record.

Past changes in surface-water nitrate use

Because phytoplankton discriminates against ^{15}N during nitrate uptake, the $\delta^{15}\text{N}$ of particulate organic matter produced in the euphotic zone increases with biological utilization of surface nitrate^{8–11} ($\delta^{15}\text{N}$ is defined in Fig. 3 legend). Isotopic balance between the new nitrate added to the euphotic zone by vertical mixing and the nitrogen removed by either biological utilization (export production) or physical processes (mixing or advection) dictates that, at steady state, the $\delta^{15}\text{N}$ of the material sinking from

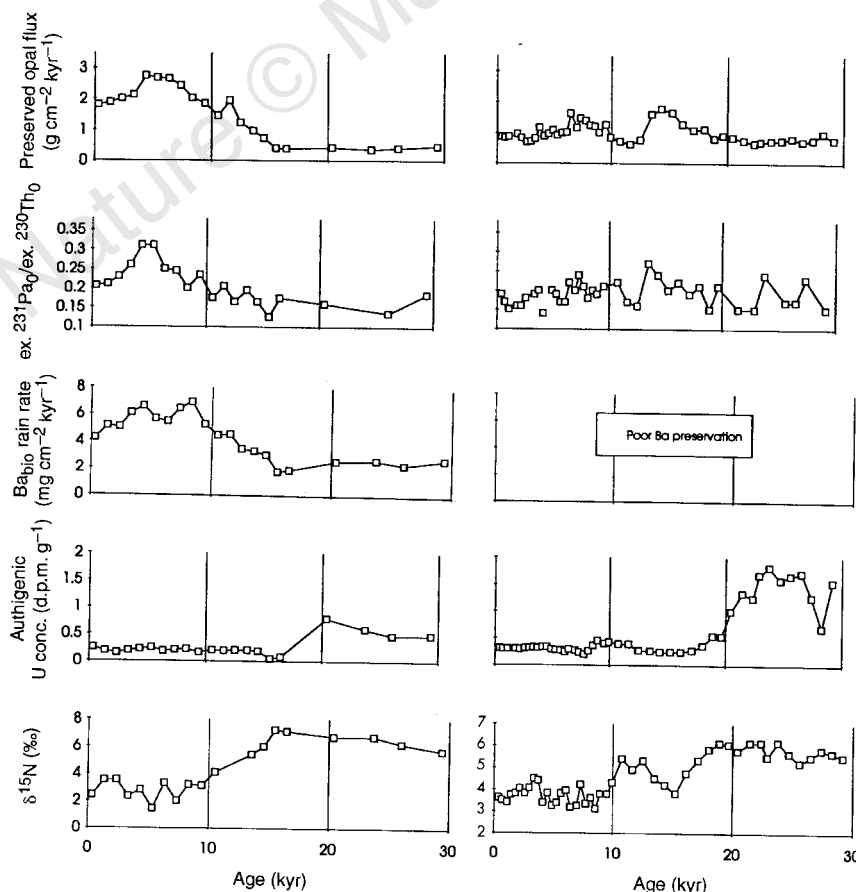


Figure 3 Age dependence of geochemical data from two sediment cores from the Indian sector of the Southern Ocean. Preserved opal flux, $\text{ex.}^{231}\text{Pa}_0/\text{ex.}^{230}\text{Th}_0$, biogenic Ba rain rate (corrected for dissolution), authigenic U concentration and bulk sediment $\delta^{15}\text{N}$ in core MD84-552 (Kerguelen plateau; $54^\circ 55' \text{S}$, $73^\circ 50' \text{E}$; 1,780 m) and MD88-773 (eastern Indian; $52^\circ 55' \text{S}$, $109^\circ 52' \text{E}$; 2,460 m). $\delta^{15}\text{N} = [({}^{15}\text{N}/{}^{14}\text{N})_{\text{sample}}/({}^{15}\text{N}/{}^{14}\text{N})_{\text{standard}}] - 1$, where $\delta^{15}\text{N}$ is in units of ‰.

the euphotic zone (δ_{sink}) is given by:

$$\delta_{\text{sink}} = \delta_{\text{NO}_3}^{\text{in}} + (\epsilon_u(1-u) \ln(1-u)/u) \quad (1)$$

where $\delta_{\text{NO}_3}^{\text{in}}$ is the $\delta^{15}\text{N}$ of the nitrate supplied to the euphotic zone (new nitrate), ϵ_u is the isotopic fractionation factor during nitrate uptake ($\epsilon_u = (k_{14}/k_{15} - 1) \times 1,000$, where k_{14} and k_{15} are the uptake rates of ^{14}N - and ^{15}N -nitrate), and u is the fraction of new nitrate that is taken up by phytoplankton. A fraction of the sinking particulate nitrogen is buried in deep-sea sediments with a relatively constant diagenetic offset (D) of 4‰, when deposited under oxic conditions (M.A.A. *et al.*, manuscript in preparation). Past changes in the $\delta^{15}\text{N}$ of sinking particles are thus recorded in bulk-sediment $\delta^{15}\text{N}$ ($\delta_{\text{sed}} = \delta_{\text{sink}} + D$), from which past changes in nitrate utilization can be estimated, if we know ϵ_u and $\delta_{\text{NO}_3}^{\text{in}}$. Because ϵ_u falls within a relatively narrow range in the modern Southern Ocean (5–6‰ south of the MPF in the region of our study^{33,34}), we assume that it stays essentially constant with time, while $\delta_{\text{NO}_3}^{\text{in}}$ can be estimated for the $\delta^{15}\text{N}$ of sediment underlying oligotrophic regions, where $u = 1$ (hence, $\delta_{\text{NO}_3}^{\text{in}} = \delta_{\text{sed}} - D$).

Reflecting the low biological utilization of surface nitrate in the Southern Ocean, $\delta^{15}\text{N}$ of Holocene sediment is invariably low south of the Polar Front (4.1‰; Fig. 4a). It gradually increases northwards, reaching a maximum north of the Subtropical Convergence (10.1‰). Taking $D = 4$ ‰, $\epsilon_u = 6$ ‰ (refs 33, 34), $\delta_{\text{NO}_3}^{\text{in}} = 5$ ‰ (refs 33, 34) and $\delta_{\text{sed}} = 4$ ‰, we calculate 30% nitrate utilization in the euphotic zone south of the Polar Front ($u = 0.3$), which is consistent with nitrate concentration profiles³⁵. North of the Polar Front, sediment $\delta^{15}\text{N}$ increases owing to a gradual increase in nitrate depletion. In this region, new nitrate appears to be supplied by the northward drift of partially depleted (and thus isotopically enriched) nitrate (Fig. 5a) originating from surface Antarctic water^{33,34}. The higher $\delta^{15}\text{N}$ of nitrate supplied to subantarctic surface water thus contributes to the higher bulk sediment $\delta^{15}\text{N}$ north of the Polar Front. The maximum in δ_{sed} just north of the Subtropical Convergence (STC) (10.1‰; Fig. 4a) is due to the cross-frontal injection of partially depleted surface nitrate by eddies crossing the front^{9–11}. In the subtropical gyre, where all surface nitrate is utilized ($u = 1$), bulk sediment $\delta^{15}\text{N}$ (8.4‰) is very close to the expected value of 9‰ (that is, $\delta_{\text{NO}_3}^{\text{in}} + D$).

During the last glacial, we find a consistently higher bulk sediment $\delta^{15}\text{N}$ south of the MPF (~ 6 ‰; Fig. 4b), indicating more efficient biological utilization of surface nitrate. This trend has now been confirmed by measuring the $\delta^{15}\text{N}$ of cleaned diatom microfossils from southern ocean cores^{34,36} (the latter results are in conflict with earlier measurements of cleaned diatom $\delta^{15}\text{N}$ which suggested much heavier Holocene values³⁷). We have carried out a number of studies in Holocene sediments which invariably confirmed the trends observed with bulk sediment $\delta^{15}\text{N}$ (ref. 36). The disagreement between the two studies is discussed elsewhere³⁴. δ_{sed} in the subtropical gyre did not change significantly between the last glacial and the Holocene, suggesting little change in the isotopic composition of the vertically supplied new nitrate. Assuming a constant D and ϵ_u , we calculate that $\sim 70\%$ of the available nitrate was utilized by phytoplankton south of the MPF during the last glacial. In contrast, in the region immediately to the north, δ_{sed} decreased slightly (Fig. 4c), producing a distinct δ_{sed} minimum (4.8‰; Fig. 4b) in the glacial ocean, centred at 45° – 50° S. Lower glacial $\delta^{15}\text{N}$ in this region would suggest a reduction in surface nitrate utilization and in the lateral supply of isotopically heavy nitrate for the south (Fig. 5b). This region appears to be much broader in the glacial Atlantic sector, where it overlaps with the region of higher glacial productivity.

Past changes in water column stratification

South of the MPF. There was a substantial increase in nitrate depletion in surface water during the last glacial (Fig. 4), and a

significant reduction of export production in the Atlantic and central Indian sector and little change in the eastern Indian sector (Fig. 1). These observations imply a significant decrease in the vertical supply rate of nitrate to the euphotic zone, and thus an increase in the degree of stratification in the upper water column. Calculating Ba_{bio} -based³¹ average Holocene export production for the Atlantic and Kerguelen plateau (Fig. 1e; $P_{\text{exp}}(\text{g C m}^{-2} \text{ yr}^{-1}) = 1.95 (\text{Ba}_{\text{bio}} \text{ rain rate})^{1.41}$, where Ba_{bio} average rain rate = $5.4 \text{ mg cm}^{-2} \text{ kyr}^{-1}$) and using Redfield stoichiometry, we estimate an export flux of particulate nitrogen $F_N = 3.5 \text{ g N m}^{-2} \text{ yr}^{-1}$ (Fig. 5a). Holocene $\delta^{15}\text{N}$ in this region averages 4‰, implying 30% nitrate utilization by phytoplankton. The supply rate of nitrate to the euphotic zone (Q_{nitrate}) is therefore $3.5 \text{ g N m}^{-2} \text{ yr}^{-1} / 0.3 = 12 \text{ g N m}^{-2} \text{ yr}^{-1}$. With a nitrate concentration at 150 m depth³⁵ ($[\text{NO}_3]_i$) of 0.4 g N m^{-3} , we calculate a water supply rate (Q_w) by vertical mixing and upwelling of $12 \text{ g N m}^{-2} \text{ yr}^{-1} / 0.4 \text{ g N m}^{-3} = 30 \text{ m yr}^{-1}$. During the last glacial (Fig. 5b), export

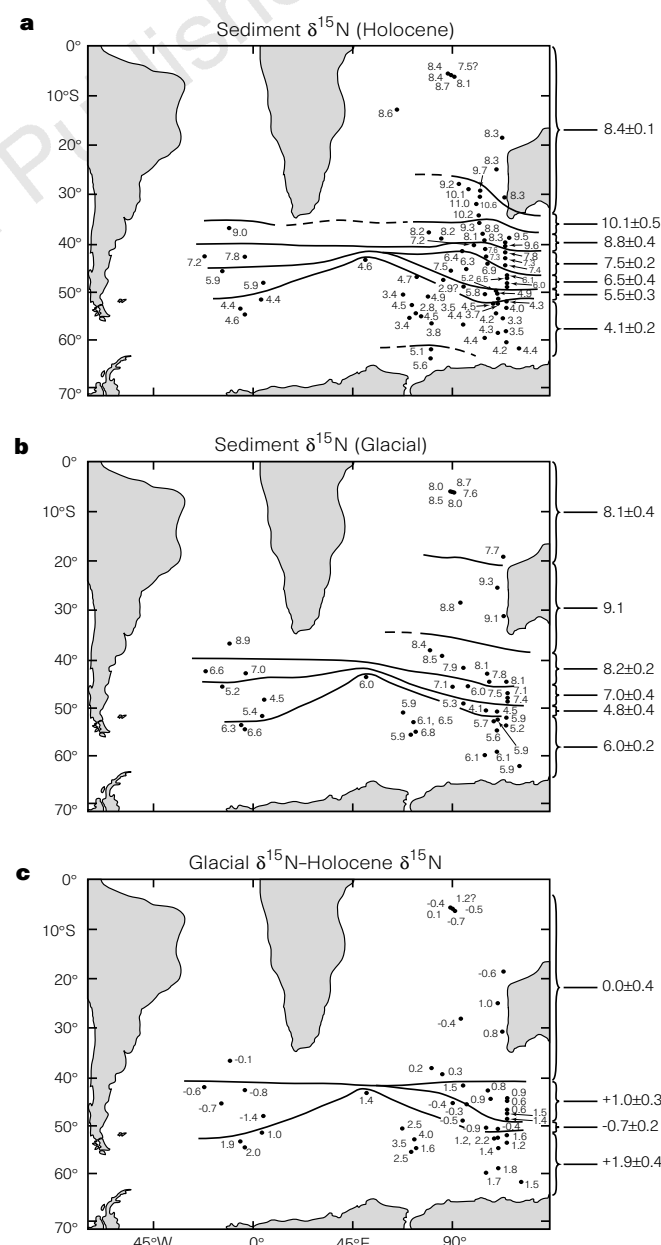


Figure 4 Bulk sediment $\delta^{15}\text{N}$. **a**, Holocene (averaged values unless single point analysis); **b**, last glacial; **c**, difference between last glacial and Holocene.

production in this region was reduced to $1.95 (2.0)^{1.41} = 5 \text{ g C m}^{-2} \text{ yr}^{-1}$ (Fig. 1f) or $0.9 \text{ g N m}^{-2} \text{ yr}^{-1}$. Glacial $\delta^{15}\text{N}$ increased to 6‰, implying 70% nitrate utilization by phytoplankton. The supply rate of nitrate to the euphotic zone must thus have been reduced to: $0.9/0.7 = 1.3 \text{ g N m}^{-2} \text{ yr}^{-1}$. If nitrate concentration at intermediate depth did not change substantially, the water supply rate during glacial would have been reduced 10-fold ($1.3 \text{ g N m}^{-2} \text{ yr}^{-1}/0.42 \text{ g N m}^{-3} = 3 \text{ m yr}^{-1}$) and nitrate concentration in surface water would have been lowered from 21 to $9 \mu\text{M}$.

Our conclusions are consistent with microfossil evidence showing the glacial predominance of *Eucampia antarctica*²⁰ and *Cyclodophora davisiana*³⁸. The predominance of the latter was taken to suggest that the hydrographic structure of the glacial Southern Ocean was similar to that of the modern Sea of Okhotsk³⁸, with a well developed near-freezing layer underlying a highly seasonal surface layer³⁹. Such structure would have limited winter mixing to the depth of 100–200 m and could account for our observations. Increased stratification is also not inconsistent with the $\delta^{13}\text{C}$ record of *Neogloboquadrina pachyderma*. *N. pachyderma* produces much of its calcite as secondary crust below surface waters⁴⁰. As a result, although the species can still record surface-water properties in the vertically mixed modern Southern Ocean⁴¹, in the stratified glacial Southern Ocean, its isotopic composition would have been heavily skewed towards that of isotopically lighter pycnocline waters.

North of the MPF. Glacial productivity increased (Atlantic sector) or did not change significantly (eastern Indian sector) whereas nitrate utilization decreased slightly, implying an increase in the rate of supply of nitrate to surface waters and increased vertical mixing. Ba-based export flux estimates for the Holocene are $7 \text{ g C m}^{-2} \text{ yr}^{-1}$ ($1.95 (2.5)^{1.41}$; Fig. 1e), or $1.2 \text{ g N m}^{-2} \text{ yr}^{-1}$ (Fig. 5a). Bulk sediment $\delta^{15}\text{N}$ is 6‰, corresponding to 60% utilization within the subantarctic region and resulting in $9 \mu\text{M}$ nitrate remaining unutilized

in surface waters. The supply rate of nitrate in the subantarctic region is thus $1.2 \text{ g N m}^{-2} \text{ yr}^{-1}/0.6 = 2 \text{ g N m}^{-2} \text{ yr}^{-1}$, and $Q_w = 2 \text{ g N m}^{-2} \text{ yr}^{-1}/0.3 \text{ g N m}^{-3} = 7 \text{ m yr}^{-1}$. During the last glacial, high sediment accumulation rates and reducing conditions prevent the use of biogenic Ba rain rate to estimate export production from all but one core (PS1754; Fig. 1f). Combining evidence from other proxies and the Ba record from this core suggests that productivity increased to a level comparable to that of the high-opal-belt region in the modern ocean ($3.5 \text{ g N m}^{-2} \text{ yr}^{-1}$). Bulk sediment $\delta^{15}\text{N}$ dropped to 5‰ (that is, 50% nitrate depletion). If nitrate concentration of the vertically supplied water remained at $\sim 30 \mu\text{M}$, the upwelling rate would have had to increase to $\sim 18 \text{ m yr}^{-1}$ and surface nitrate concentration to $15 \mu\text{M}$ (Fig. 5b).

Implication for atmospheric CO₂ and ocean circulation

Earlier box models^{1–6} have suggested two primary high-latitude ocean mechanisms that could reduce atmospheric CO₂ during glacial periods: higher productivity or lower vertical mixing. Our data support decreased vertical mixing in the water column south of the MPF as a mechanism for lowering atmospheric CO₂. North of the MPF, increased productivity was also accompanied by increased vertical mixing. The rate of supply of nutrients and CO₂ to the euphotic zone of this region would thus have been higher, counteracting the effect of stratification farther south. The combined effect of these changes in vertical mixing rates, however, seems to have resulted in an overall increase in the mean surface nitrate utilization. If the zone of increased stratification was about twice the area of the zone of increased mixing (Fig. 4), and assuming a similar trend in the Pacific sector, the mean surface nitrate utilization would have decreased from 0.63 ($(0.33 \times 0.5) + (0.66 \times 0.7)$) during glacial times to 0.43 during the Holocene ($(0.33 \times 0.7) + (0.66 \times 0.3)$). The net effect would have thus contributed to the lowering of glacial atmospheric CO₂. Preliminary calculations performed with a modified CYCLOPS box model^{34,42} suggest a decrease which could account for $\sim 50\%$ of the glacial lowering in atmospheric CO₂ ($\sim 40 \text{ p.p.m.}$).

The modified CYCLOPS box-model also predicts a decrease in the oxygen concentration of deep Southern Ocean waters to $85 \mu\text{M}$, consistent with our interpretation of the authigenic U profiles. A decrease in the deep-water concentration of O₂ is a direct consequence of increased nitrate depletion in surface high-latitude water^{1–6,42,43}. Evidence from the distribution of ²³¹Pa/²³⁰Th ratio in sediments²⁷ and $\delta^{13}\text{C}$ and Cd/Ca in benthic foraminifera⁴⁴ indicates that the Glacial Upper Circumpolar Deep Water was well ventilated during the last glacial by North Atlantic Deep/Intermediate Water (GNAIW). As a result, reduced oxygen concentrations induced by water-column stratification in the southernmost region affected mainly the lower Circumpolar Deep Water, which appears to have been largely isolated from the glacial global conveyor-belt circulation^{44,45}. A lack, or reduction, of Antarctic Bottom Water formation as a result of grounded ice in areas of formation of this water mass⁴⁶ could also have contributed to reducing ventilation in bottom Antarctic waters⁴⁵. This model can also be reconciled with the benthic foraminifera Cd/Ca record in the Southern Ocean. Little change in the Cd/Ca ratio is expected at intermediate depths ventilated by GNAIW, whereas in deeper water, the expected glacial increase in Cd/Ca could have been reduced by increased carbonate dissolution⁴⁷ resulting from a shoaling of the lysocline⁴⁸. □

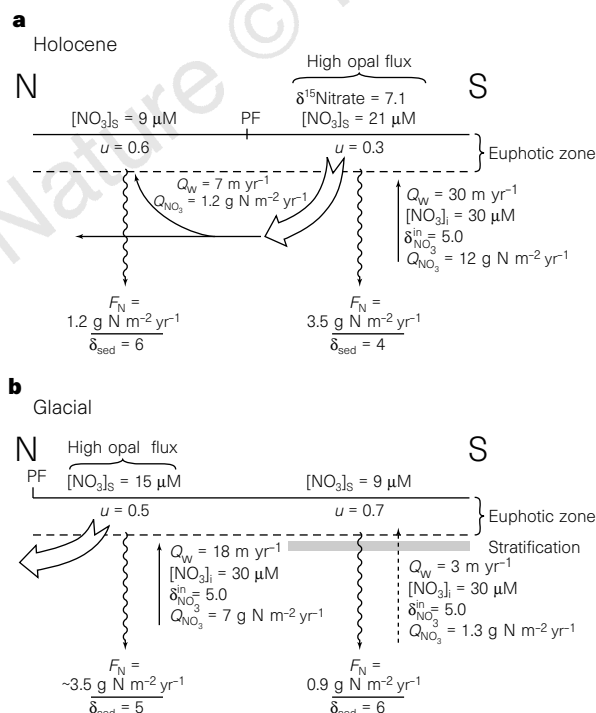


Figure 5 Changes in export flux, vertical mixing and nutrient utilization inferred for the Atlantic and central Indian sector of the Southern Ocean. **a**, Holocene; **b**, last glacial (see text for explanations).

Received 12 September 1996; accepted 3 September 1997.

1. Sarmiento, J. L. & Toggweiler, R. A new model for the role of the oceans in determining atmospheric pCO₂. *Nature* **308**, 621–624 (1984).
2. Siegenthaler, U. & Wenk, T. Rapid atmospheric CO₂ variations and ocean circulation. *Nature* **308**, 624–626 (1984).
3. Knox, F. & McElroy, M. B. Changes in atmospheric CO₂: Influence of the marine biota at high latitude. *J. Geophys. Res.* **89**, 4629–4637 (1984).
4. Toggweiler, R. & Sarmiento, J. L. in *The Carbon Cycle and Atmospheric CO₂: Natural Variations Archean to Present* (eds Sundquist, E. T. & Broecker, W. S.) 163–184 (Vol. 32, Geophys. Monogr. Ser., Am. Geophys. Union, Washington DC, 1985).

5. Wenk, T. & Siegenthaler, U. in *The Carbon Cycle and Atmospheric CO₂: Natural Variations Archaean to Present* (eds Sundquist, E. T. & Broecker, W. S.) 185–194 (Vol. 32, Geophys. Monogr. Ser., Am. Geophys. Union, Washington DC, 1985).
6. Knox-Ennever, F. & McElroy, M. B. in *The Carbon Cycle and Atmospheric CO₂: Natural Variations Archaean to Present* (eds Sundquist, E. T. & Broecker, W. S.) 154–162 (Vol. 32, Geophys. Monogr. Ser., Am. Geophys. Union, Washington DC, 1985).
7. Kumar, N. *et al.* Increased biological productivity and export production in the glacial southern ocean. *Nature* **378**, 675–680 (1995).
8. Altabet, M. A. *et al.* Seasonal and depth-related changes in the source of sinking particles in the North Atlantic detected using ¹⁵N/¹⁴N ratios. *Nature* **354**, 136–139 (1991).
9. Francois, R., Altabet, M. A. & Burckle, L. D. Glacial to interglacial changes in surface nitrate utilization in the Indian sector of the southern ocean as recorded by sediment $\delta^{15}\text{N}$. *Paleoceanography* **7**, 589–606 (1992).
10. Altabet, M. A. & Francois, R. Sedimentary nitrogen isotopic ratio as a recorder for surface ocean nitrate utilization. *Glob. Biogeochem. Cycles* **8**, 103–116 (1994).
11. Altabet, M. A. & Francois, R. in *Carbon Cycling in the Glacial Ocean: Constraints on the Ocean's Role in Global Change* (eds Zahn, R., Pedersen, T. F., Kaminski, M. & Labeyrie, L. D.) 281–306 (Vol. 17, NATO AST Ser. I, Springer, Berlin, 1994).
12. Mortlock, R. A. *et al.* Evidence for lower productivity in the Antarctic Ocean during the last glaciation. *Nature* **351**, 220–223 (1991).
13. Charles, C. D. *et al.* Biogenic opal in southern ocean sediments over the last 450,000 years: Implications for surface water chemistry and circulation. *Paleoceanography* **6**, 697–728 (1991).
14. Bareille, G. *et al.* Biogenic silica accumulation rate during the Holocene in the southeastern Indian Ocean. *Mar. Chem.* **35**, 537–551 (1991).
15. Suman, D. O. & Bacon, M. P. Variations in Holocene sedimentation in the North American basin determined from ²³⁰Th measurements. *Deep-Sea Res.* **36**, 869–878 (1989).
16. Francois, R., Bacon, M. P. & Suman, D. O. Thorium-230 profiling in deep-sea sediments: High-resolution records of flux and dissolution of carbonate in the equatorial Atlantic during the last 24,000 years. *Paleoceanography* **5**, 761–787 (1990).
17. Kumar, N. *et al.* ²³¹Pa/²³⁰Th ratios in sediments as a proxy for the past changes in southern ocean productivity. *Nature* **362**, 45–48 (1993).
18. van Bennekom, A. J. *et al.* Primary productivity and the silica cycle in the southern ocean (Atlantic sector). *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **67**, 19–30 (1988).
19. Frank, M. Reconstruction of late Quaternary environmental conditions applying the natural radionuclides ²³⁰Th, ¹⁰Be, ²³¹Pa and ²³⁸U: A study of deep-sea sediments from the eastern sector of the Antarctic Circumpolar Current system. Thesis, Heidelberg Univ. (1995).
20. Burckle, L. H. Diatom distribution and paleoceanographic reconstruction in the southern ocean: Implications for late Quaternary paleoceanography. *Mar. Micropaleontol.* **9**, 241–261 (1984).
21. Morley, J. J. Variations in high-latitude oceanographic fronts in the southern Indian Ocean: An estimation based on faunal changes. *Paleoceanography* **4**, 547–554 (1989).
22. Sullivan, C. W. *et al.* Distribution of phytoplankton blooms in the southern ocean. *Science* **262**, 1832–1837 (1993).
23. Bacon, M. P. Tracers of chemical scavenging in the ocean: boundary effects and large scale chemical fractionation. *Phil. Trans. R. Soc. Lond. A* **320**, 187–200 (1988).
24. Anderson, R. F. *et al.* Boundary scavenging in the Pacific Ocean: a comparison of ¹⁰Be and ²³¹Pa. *Earth Planet. Sci. Lett.* **96**, 287–304 (1990).
25. Yu, E.-F. Variations in the particulate flux of ²³⁰Th and ²³¹Pa and paleoceanographic applications of the ²³¹Pa/²³⁰Th ratio. Thesis, WHOI-MIT (1994).
26. Walter, H.-J., Rutgers van der Loeff, M. M. & Hoeltzen, H. Enhanced scavenging of ²³¹Pa relative to ²³⁰Th in the south Atlantic south of the Polar Front: Implications for the use of the ²³¹Pa/²³⁰Th ratio as a paleoproductivity proxy. *Earth Planet. Sci. Lett.* **149**, 85–100 (1997).
27. Yu, E.-F., Francois, R. & Bacon, M. P. Similar rates of modern and last-glacial ocean thermohaline circulation inferred from radiochemical data. *Nature* **379**, 689–694 (1996).
28. Klinkhammer, G. & Palmer, M. R. Uranium in the oceans, where it goes and why. *Geochim. Cosmochim. Acta* **55**, 1799–1806 (1991).
29. Bareille, G. *et al.* Glacial-Interglacial changes in the accumulation rates of major biogenic components in southern Indian Ocean sediments. *J. Mar. Syst.* (in the press).
30. Dymond, J., Suess, E. & Lyle, M. Barium in deep-sea sediment: A geochemical indicator of paleoproductivity. *Paleoceanography* **7**, 163–181 (1992).
31. Francois, R. *et al.* Biogenic Ba fluxes to the deep-sea: Implications for the paleoproductivity reconstruction. *Glob. Biogeochem. Cycles* **9**, 289–303 (1995).
32. Nurnberg, C. C. Bariumfluss und sedimentation im sudlichen Sudatlantik-Hinweise auf produktivitätsänderungen im Quartar. Thesis, Kiel Univ. (1995).
33. Sigman, D. M. *et al.* in *ASLO, Aquatic Sci. Meeting Abstr.* Santa Fe, New Mexico, 304 (1997).
34. Sigman, D. M. The role of biological production in Pleistocene atmospheric carbon dioxide variations and the nitrogen isotope dynamics of the Southern Ocean. Thesis, WHOI-MIT (1997).
35. GEOSECS *Sections and Profiles* (National Science Foundation, Washington DC, 1982).
36. Sigman, D. M. *et al.* Diatom microfossil nitrogen isotope composition supports the hypothesis of higher nitrate utilization in the glacial southern ocean. *Eos* **78**, S190 (1997).
37. Shemesh, A. *et al.* Isotopic evidence for reduced productivity in the glacial southern ocean. *Science* **262**, 407–410 (1993).
38. Morley, J. J. & Hays, J. D. Oceanographic conditions associated with the high abundance of radiolarian *Cyclodophora davisiana*. *Earth Planet. Sci. Lett.* **66**, 63–72 (1983).
39. Yang, J. & Honjo, S. J. Modeling the near-freezing dichothermal layer in the sea of Okhotsk and its interannual variations. *J. Geophys. Res.* **101**, 16421–16433 (1996).
40. Kohfeld, K. E., Fairbanks, R. G., Smith, S. L., Walsh, I. D. Neoglobobulimina pachyderma (sinistral coiling) as paleoceanographic tracers in polar oceans: Evidence from Northeast Water Polynya plankton tows, sediment traps and surface sediments. *Paleoceanography* **11**, 679–700 (1996).
41. Charles, C. D. & Fairbanks, R. G. in *Geological History of the Polar Oceans: Arctic versus Antarctic* (eds Bleil, U. & Thiede, J.) 519–538 (Kluwer Academic, Norwell, MA, 1990).
42. Sigman, D. & McCorkle, D. Comparing the closed and open system effects of changes in low latitude biological production using a reservoir model of the ocean carbon cycle. *Eos* **75**, 367 (1994).
43. Keir, R. S. On the late Pleistocene ocean geochemistry and circulation. *Paleoceanography* **3**, 413–445 (1988).
44. Lynch-Stieglitz, J., van Geen, A. & Fairbanks, R. G. Inter-ocean exchange of Glacial North Atlantic Intermediate Water: Evidence from Subantarctic Cd/Ca and carbon isotope measurements. *Paleoceanography* **11**, 191–202 (1996).
45. Michel, E. *et al.* Could deep Subantarctic convection feed the world deep basins during the last glacial maximum? *Paleoceanography* **10**, 927–942 (1995).
46. Kellogg, T. B. Glacial-Interglacial changes in global deepwater circulation. *Paleoceanography* **2**, 259–271 (1987).
47. McCorkle, D. C. *et al.* Evidence of a dissolution effect on benthic foraminiferal shell chemistry: $\delta^{13}\text{C}$, Cd/Ca, Ba/Ca, and Sr/Ca results from the Ontong Java Plateau. *Paleoceanography* **10**, 699–714 (1995).
48. Howard, W. R. & Prell, W. L. Late Quaternary CaCO₃ production and preservation in the Southern Ocean: Implications for oceanic and atmospheric carbon cycling. *Paleoceanography* **9**, 453–482 (1994).
49. Bareille, G. Flux sédimentaires: paléproductivité et paléocirculation de l'Océan Austral au cours des 150,000 dernières années. Thesis, Univ. Bordeaux (1994).

Acknowledgements. We thank A. P. Fleer for assistance with radiochemical analysis, and R. F. Anderson, W. R. Howard, M. M. Rutgers van der Loeff and H.-J. Walter for reviews and criticisms. This work was supported by NSF and also benefited from the use of the MIT research reactor with support from a US DOE Reactor Sharing Grant.

Correspondence and requests for materials should be addressed to R.F. (e-mail: rfrancois@whoi.edu).

KNOW YOUR COPYRIGHTS RESPECT OURS

The publication you are reading is protected by copyright law.
Photocopying copyright material without permission is no different from stealing a magazine from a newsagent, only it doesn't seem like theft.

If you take photocopies from books, magazines and periodicals at work your employer should be licensed with CLA.
Make sure you are protected by a photocopying licence.



The Copyright Licensing Agency Limited
90 Tottenham Court Road, London W1P 0LP
Telephone: 0171 436 5931 Fax: 0171 436 3986

Quantum oscillations in a confined electron gas

Ch. Würsch*, C. Stamm*, S. Egger*, D. Pescia*, W. Baltensperger† & J. S. Helman†‡

* Laboratorium für Festkörperphysik der ETH Zürich, CH-8093 Zürich, Switzerland

† Centro Brasileiro de Pesquisas Físicas (CBPF), Rua Dr Xavier Sigaud 150, Rio de Janeiro, RJ 22290, Brazil

When metals are structured on nanometre length scales, their electrons are subject to confinement effects: the response of a confined electron gas is governed by Friedel oscillations¹ of the electron density and Rudermann–Kittel–Kasuya–Yosida oscillations² of the spin density. Spatial oscillations of electron density have been observed directly at surfaces (in the vicinity of defects and steps) by scanning tunnelling spectroscopy^{3,4}. But it has proved more difficult to probe such oscillations in bulk materials and over large distances⁵. Here we report the detection of quantum oscillations in a three-dimensional electron gas confined to a half space by a surface. To facilitate this detection, we have inserted an atomically thin ferromagnetic cobalt film at a variable distance τ from the surface of a copper single crystal. The cobalt film induces⁵ a total spin polarization P in the conduction electrons of the copper and, by virtue of the confining effects of the copper–vacuum interface, P varies as a function of τ . Our measurements of P reveal both quantum oscillations (the wavelengths of which are governed by the extremal diameters of the copper Fermi surface) and a decay with τ that are consistent with theoretical expectations². These observations show that a consequence of improving the quality of nanostructured materials is that long-range quantum interactions can emerge more effectively, so that even distant boundaries and defects can become pivotal in determining physical properties.

To describe the quantum oscillations we consider a degenerate free-electron gas in a half space and a point-like spin-dependent coupling to the electrons at a distance τ from the boundary² of the half space (Fig. 1). As in the theory of Rudermann–Kittel–Kasuya–Yosida (RKKY), the magnetic impurity perturbing the electron gas is represented by the hamiltonian² $-\gamma\delta(x-\tau)\delta^2(\mathbf{R})\sigma_z/2$ where $\mathbf{R} = \{y, z\}$ varies over the plane parallel to the surface, δ is the Dirac delta function, and γ represents the exchange coupling between the spin $\frac{1}{2}\sigma_z$ of the conduction electrons and point-like perturbations. This magnetic coupling is the source of the spatially oscillating RKKY spin density $\sigma(\mathbf{r})$ in the conduction electron gas. The quantity we are interested in is the integral P of $\sigma(\mathbf{r})$ over $\mathbf{r}$, that is, total spin polarization of the semi-infinite electron gas. A theorem² relates P to the unperturbed energy-resolved electron density $\rho(\mathbf{r}, E_F)$ with $\mathbf{r}$ taken at the site of the impurity. Here E_F is the Fermi energy of the electron gas, and $\rho(\mathbf{r}, E)$ is the local electron density at a well-defined energy E (ref. 6), and is, for example, the quantity measured by scanning tunnelling spectroscopy^{3,4}. When electrons are confined to a half space by a surface, the energy-resolved electron density is expected to display quantum oscillations as a function of the distance τ from the surface⁶. The wavelength of the oscillation of $\rho(\tau, E_F)$ is determined by extremal diameters of the Fermi surface⁷. Its amplitude decays, for a two-dimensional surface confining a three-dimensional electron gas, as the inverse of the distance τ (refs 2, 6).

Locating a magnetic impurity at any site inside a plane lying at a distance τ from the surface (see Fig. 1 inset) induces a total spin polarization $P(\tau)$ in the confined electron gas reproducing exactly

the quantum oscillations of the energy-resolved electron density of the semi-infinite electron gas. Our experiment deals with a planar structure of magnetic impurities parallel to the surface. As each magnetic impurity contributes the same $P(\tau)$, the total spin polarization resulting from a planar array has the same functional dependence on τ as the total spin polarization originating from one magnetic impurity. Fig. 1 shows the calculated $P(\tau)$ (continuous line) in units of the bulk total spin polarization $P_0 = k_F\gamma m/(4\pi^2\hbar)$. For the calculation in Fig. 1 we have assumed, for simplicity, that the electron gas has a spherical Fermi surface and we have represented the surface by an infinite potential step. Thus, the wavelength of the oscillation, given solely by the radius k_F of the Fermi surface, is π/k_F . For a comparison, we also show in Fig. 1 the behaviour of the total electron density (dashed curve), obtained by integrating the energy-resolved electron density over all occupied energy states. Although oscillating with the same wavelength, this last quantity, as opposed to P , decays as $1/\tau^2$ (ref. 1). We note that the system described here might have quantum-well states, the number of which changes when τ increases by π/k_F . It has been shown^{8–10} that these bound states do not produce an oscillatory term in the total energy of the system.

The samples used in our experiments consist of atomically thin films of ferromagnetic Co deposited by molecular beam epitaxy on top of a Cu(100) (ref. 11) single crystal, slightly miscut to induce a preferential step direction. The Co film is covered by a Cu film with variable thickness τ (Fig. 2a). The metallic overlayers largely reproduce the preferential step direction¹¹. Thus, the Co film has a two-fold easy magnetization axis^{12,13} along the step direction. Applying an in-plane magnetic field H perpendicular to the easy axis results in shifted M versus H hysteresis loops (Fig. 2a), M being the total magnetization of the multilayer structure as seen by the magneto-optic Kerr effect^{12,13}. The shifting of the loop is due to the total magnetization switching abruptly (except for a small hysteresis) from the easy direction into the direction specified by the external magnetic field when H reaches a threshold value H_s . The value of H_s is the result of the two-fold step induced anisotropy balanced against the four-fold anisotropy and the magnetic field energy to minimize the total energy of the system (C.W. and D.P.,

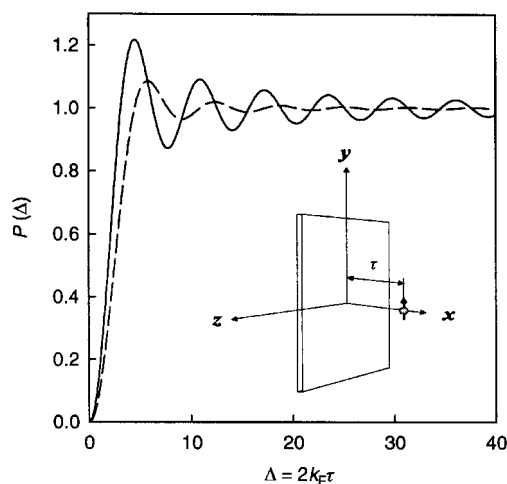


Figure 1 Quantum oscillations: theory. Inset shows the definition of τ . The graph shows the total spin polarization $P(\Delta)$ (where $\Delta = 2k_F\tau$) of the conduction electrons in a semi-infinite free-electron gas, in units of the asymptotic value $P_0 = P(\Delta \rightarrow \infty)$. For an infinite step potential simulating the surface, we can obtain the exact expression $P(\Delta)/P_0 = 1 - [\sin \Delta]/\Delta$, given as the solid line. P vanishes for $\Delta = 0$, as in this model there are no electrons at the surface of the half space. $P(\Delta)$ rises within the distance $\Delta \approx \pi$ to the value corresponding to the Pauli spin polarization P_0 . Then it oscillates with a wavelength 2π , the amplitude of the oscillation decaying as $1/\Delta$. Also shown (dashed line) is the total electron density of an ideal Fermi gas in half space, in units of the density at large distances.

‡ Deceased.

manuscript in preparation). Part of the magnetic field energy is given by the total spin polarization $P(\tau)$ of the Cu sp -electrons coupling to the external magnetic field. Through this part, H_s acquires the oscillatory component of $P(\tau)$. A further source of the oscillatory component is, in principle, the spatially varying $\rho(\mathbf{r}, E_F)$ affecting both the value of the magnetic anisotropies and the magnetic moments of the buried Co film. As the bulk total spin polarization P_0 of the Cu conduction electrons is estimated⁵ to be as much as 10% of the impurity moment, we expect $P(\tau)$ to be the most important source of the observed quantum oscillations. The behaviour of H_s versus τ is shown in Fig. 2b. After a few strong oscillations at low coverages¹², the asymptotic ($\tau \geq 6$ monolayers (ML)) part of $H_s(\tau)$ (suitably rescaled in Fig. 2b) shows a slowly decaying oscillatory component, which contains essentially two wavelengths as revealed by the Fourier transform: a dominating short wavelength $\lambda \approx 2.4$ ML and a longer wavelength $\Lambda \approx 5.4$ ML. These wavelengths are determined by the two extremal diameters of the Cu Fermi surface along the (100) direction⁷. The stronger weight of λ is in line with theoretical expectations for samples with perfect geometry⁷. The decay of the amplitude of the oscillations is best

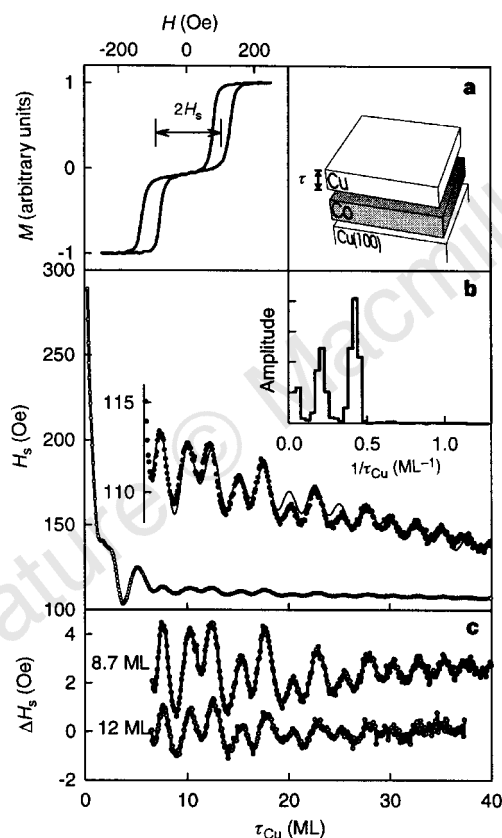


Figure 2 Quantum oscillations: experiment. **a**, The left-hand side shows a typical shifted hysteresis loop M versus H for a system (schematically shown on the right-hand side) consisting of a Co film inserted at a variable distance τ from the surface of a semi-infinite Cu crystal. H is applied in the film plane and perpendicularly to the easy magnetization direction. At $H = H_s$ the magnetization switches via a first-order phase transition (accompanied by a small hysteresis) into the direction specified by the magnetic field. **b**, H_s versus τ , plotted with two different vertical scales. Inset, the Fourier transform of the asymptotic part of the experimental data ($\tau > 6$ ML). Simultaneous Cu deposition and measurement were performed at ~ 170 K. The two main wavelengths derived from the Fourier transform are used to fit the asymptotic part of the experimental data (see the continuous curve). The decay of the oscillatory part is best fitted by a $1/\tau$ power-law dependence. **c**, H_s (background subtracted) versus τ for samples with different Co thicknesses.

fitted by a $1/\tau$ power law (the continuous curve in Fig. 2b). Varying the thickness of the Co film leaves the oscillatory component essentially unchanged (Fig. 2c).

The observation of the quantum oscillations requires τ to be well defined over large lateral distances. We find that depositing the Cu capping layer at room temperature produces the 'pyramid-like' morphology shown in Fig. 3a. Accordingly, structures grown at room temperature do not¹³ display the well-defined, long-ranged oscillatory component shown in Fig. 2b. Instead, growing the epitaxial structures at low temperature, as in Fig. 2b, avoids the formation of 'pyramids', thereby reducing thickness fluctuations (Fig. 3b). The morphological findings of Fig. 3 are confirmed by the magnetic results (Fig. 4) on the coupling between two Co-films separated by a Cu wedge¹⁴. The wedge-like structure, schematically given in Fig. 4a, provides a variable Cu interlayer thickness τ . The quantity measured as a function of τ is the coupling field H_j , recorded by measuring the total magnetization M of the system versus the externally applied magnetic field H by means of the magneto-optical Kerr effect¹⁴ (Fig. 4a). H_j is the field required to align the magnetization of antiferromagnetically ordered films into

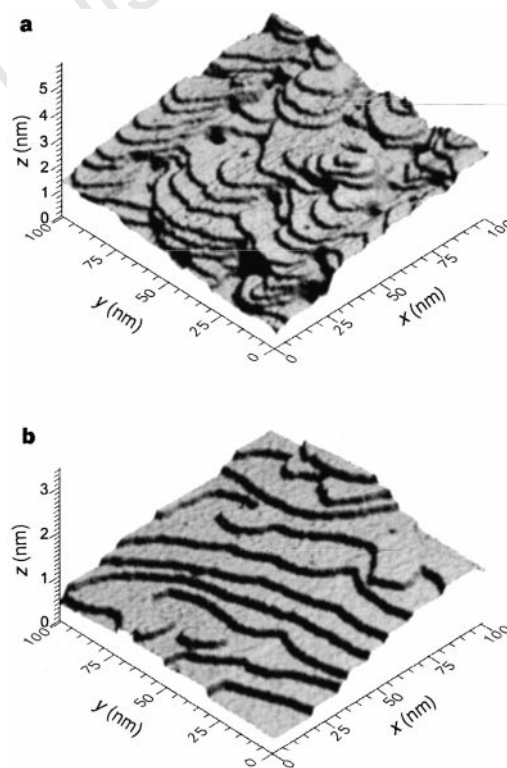


Figure 3 Effect of substrate temperature on the morphology of the deposited films. **a**, Room-temperature scanning tunnelling microscopy (STM) topography of 17.5-ML Cu film deposited by molecular beam epitaxy on top of a 9.5-ML Co film. The area covered by the image is 100×100 nm. The two films were grown successively on top of a Cu(100) single crystal, slightly miscut (0.7°) to induce monoatomic surface steps along a preferential direction. The miscut gives rise to the overall slope present in the STM images. All films were grown at ~ 320 K. The height variation superimposed on the slope is 1.5 nm. Because of the 'pyramid-like' corrugation as many as ~ 10 Cu planes are visible by a nominal thickness of 17.5 ML. **b**, Room-temperature STM topography (100×100 nm) of 14.6-ML-thick Cu film grown on top of an 8.6-ML-thick Co film, the substrate being the same as in **a**. In contrast to **a**, the Cu film was grown at ~ 150 K. The film roughly reproduces the substrate-induced monoatomic steps, and no 'pyramids' are detected.

the same direction. A zero value of H_j indicates ferromagnetically coupled films. H_j is known^{7,14} to oscillate as a function of τ because of the RKKY interaction. We can clearly distinguish two situations. In Fig. 4b (Cu spacer grown at room temperature) only a few oscillations take place, the oscillation with the short wavelength λ being marginal. This is in contrast to the theoretical expectations⁷, which require λ to dominate over Λ for geometrically perfect samples. On the other hand, when the Cu spacer is grown at low temperatures, more oscillations appear (Fig. 4c), and λ dominates over Λ . We note that cooling a room-temperature-deposited wedge structure does not change the results given in Fig. 4b. Thus, low-temperature deposition is essential to minimize thickness fluctuations. We also note that, in for example Fe/Cr multilayers, just the opposite occurs: the shortest-wavelength oscillation is favoured by high-temperature growth¹⁵.

The smallest oscillations detected in this experiment correspond to an energy change of $\sim 10^{-8}$ eV per atom. This value represents a challenge both to first-principles calculations and to spectroscopies aimed at unveiling the electronic structure of solids. Moreover, although exceedingly small, this energy scale participates in deter-

mining a macroscopic observable (in this case the threshold magnetic field H_s) and thus might play an important role in the design of metallic nanostructures for device application. □

Received 26 February; accepted 19 August 1997.

- Lang, N. D. & Kohn, W. Theory of metal surfaces: charge density and surface energy. *Phys. Rev. B* **1**, 4555–4568 (1970).
- Helman, J. S. & Baltensperger, W. Rudermann–Kittel–Kasuya–Yosida polarization in some inhomogeneous situations. *Phys. Rev. B* **51**, 12682–12691 (1994).
- Crommie, M. F., Lutz, C. P. & Eigler, D. M. Imaging standing waves in a two-dimensional electron gas. *Nature* **363**, 524–527 (1993).
- Hasegawa, Y. & Avouris, Ph. Direct observations of standing wave formation at surface steps using scanning tunneling spectroscopy. *Phys. Rev. Lett.* **71**, 1071–1074 (1993).
- Boyce, J. B. & Slichter, C. P. Conduction-electron spin density around Fe impurities in Cu above and below the Kondo temperature. *Phys. Rev.* **13**, 379–396 (1976).
- Werner, C., Schulte, F. K. & Bross, H. Surface densities of state in the planar uniform background model. *J. Phys. C* **8**, 3817–3824 (1975).
- Bruno, P. & Chappert, C. Oscillatory coupling between ferromagnetic layers separated by a nonmagnetic metal spacer. *Phys. Rev. Lett.* **67**, 1602–1605 (1991).
- Bruno, P. Comment on “Contribution of quantum-well states to the RKKY coupling in magnetic multilayers”. *Phys. Rev. Lett.* **72**, 3627–3628 (1994).
- Yosida, K. & Okiji, A. Long-range magnetic coupling in metals. *Phys. Rev. Lett.* **14**, 301–302 (1965).
- Bardasis, A. *et al.* Possibility of long-range spin polarisation in a degenerate electron gas. *Phys. Rev. Lett.* **14**, 298–302 (1965).
- Ramsperger, U., Vaterlaus, A., Pfäffli, P., Maier, U. & Pescia, D. Growth of Co on a stepped and on a flat Cu(100) surface. *Phys. Rev. B* **52**, 8001–8006 (1996).
- Weber, W., Back, C. H., Bischof, A., Pescia, D. & Allenspach, R. Magnetic switching in cobalt films by adsorption of copper. *Nature* **374**, 788–791 (1995).
- Weber, W. *et al.* Oscillatory magnetic anisotropy and quantum well states in Cu/Co/Cu(100) films. *Phys. Rev. Lett.* **76**, 3424–3427 (1996).
- Qiu, Z., Pearson, J. & Bader, S. D. Oscillatory interlayer magnetic coupling of wedged Co/Cu/Co sandwiches grown on Cu(100) by molecular beam epitaxy. *Phys. Rev. B* **46**, 8659–8662 (1992).
- Unguris, J., Celotta, R. J. & Pierce, D. T. Observation of two different oscillation periods in the exchange coupling of Fe/Cr/Fe(100). *Phys. Rev. Lett.* **67**, 140–143 (1991).

Acknowledgements. We thank R. Monnier and G. Rossi for discussions. This work was supported by the Schweizerische Nationalfonds zur Förderung der wissenschaftliche Forschung.

Correspondence and requests for materials should be addressed to D.P. (e-mail: pescia@solid.phys.ethz.ch).

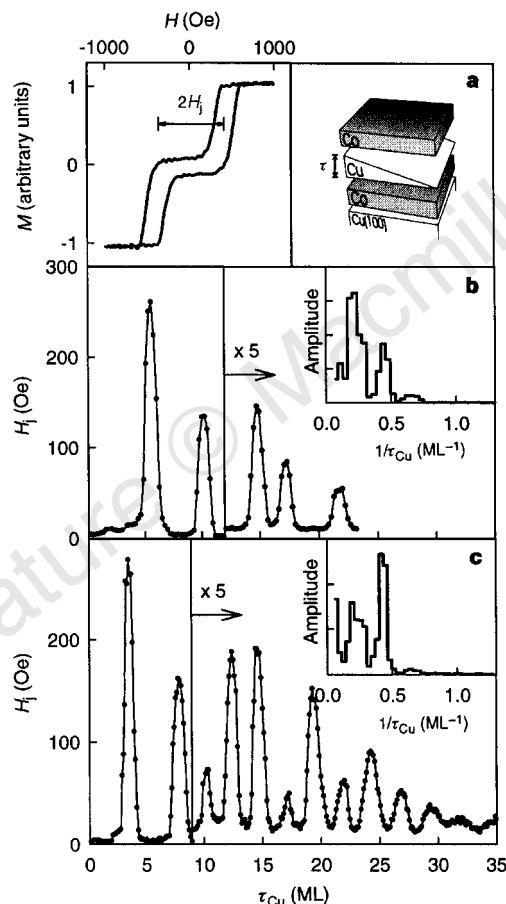


Figure 4 Enhancement of the short-wavelength component of the interlayer coupling by low-temperature deposition. **a**, The left-hand side shows a typical hysteresis curve (M versus H) recorded for exchange-coupled Co films. At $H = H_j$ the magnetizations of the individual films are aligned to the direction specified by the external magnetic field. The critical field H_j is measured as function of the Cu spacer thickness τ by scanning a focused laser beam over a wedge-like multilayered structure, shown schematically on the right-hand side. **b**, H_j versus τ for a room-temperature-grown wedge-like multilayered structure. The thicknesses of the Co films are 13.2 ML and 15.8 ML, respectively. Inset, the Fourier transform, the two peaks corresponding to wavelengths of ~ 2.4 ML and ~ 5.4 ML. The long wavelength dominates. **c**, As **b** but with the Cu wedge and the final Co film deposited and measured at ~ 160 K. The short wavelength now dominates (see Fourier transforms, inset).

P-type electrical conduction in transparent thin films of CuAlO_2

Hiroshi Kawazoe, Masahiro Yasukawa*, Hiroyuki Hyodo, Masaaki Kurita, Hiroshi Yanagi & Hideo Hosono

Materials and Structures Laboratory, Tokyo Institute of Technology, Nagatsuta, Midori-ku, Yokohama 226, Japan

Optically transparent oxides tend to be electrical insulators, by virtue of their large electronic bandgap (≥ 3.1 eV). The most notable exceptions are doped versions of the oxides In_2O_3 , SnO_2 and ZnO —all n-type (electron) conductors—which are widely used as the transparent electrodes in flat-panel displays^{1,2}. On the other hand, no transparent oxide exhibiting high p-type (hole) conductivity is known to exist, whereas such materials could open the way to a range of novel applications. For example, a combination of the two types of transparent conductor in the form of a pn junction could lead to a ‘functional’ window that transmits visible light yet generates electricity in response to the absorption of ultraviolet photons. Here we describe a strategy for identifying oxide materials that should combine p-type conductivity with good optical transparency. We illustrate the potential of this approach by reporting the properties of thin films of CuAlO_2 , a transparent oxide having room-temperature p-type conductivity up to 1 S cm^{-1} . Although the conductivity of our candidate material is significantly lower than that observed for the best n-type conducting oxides, it is sufficient for some applications, and demonstrates that the development of transparent p-type conductors is not an insurmountable goal.

The recent discoveries of several new transparent conducting

* Present address: National Industrial Research Institute of Nagoya, 1-1 Hirate-cho, Kita-ku, Nagoya 462, Japan.

oxides of n-type (see, for example, refs 3–5), and trials^{6,7} to fabricate field-effect transistors from the materials, are in striking contrast with the lack of reports of p-type conducting oxides. The non-existence of p-type transparent conducting oxides is thought to originate from a general characteristic in the electronic structure of oxides: the strong localization of the upper edge of the valence band to oxide ions. Therefore, any finding of a p-type conducting oxide must include modification of the energy band structure to reduce the localization behaviour, which in turn requires new insight into the relation between electronic structure and properties of oxide materials. In the technological field, finding such a material may open the way to new applications such as ultraviolet-emitting diodes.

First, let us consider the requirements for the major cationic species of candidate oxides if the localization behaviour in the valence band of typical oxides is to be modified. The cationic species is required to have a closed shell whose energy is almost comparable to those of the $2p$ levels of oxygen anions. The closed shell valence state is required to avoid coloration due to intra-atomic excitations. The cations we selected are Cu^+ , Ag^+ and Au^+ , which have the electronic configuration $d^{10}s^0$. Within the group of cations having the same electronic configuration, the energy of the d^{10} closed shell electrons is highest for these three cations, and is expected to overlap with that of the $2p$ electrons on oxide ions.

The second condition to be considered in the selection of the candidate oxide is the crystal structure which enhances the covalency in the bonding between the cation and oxide ion. Tetrahedral coordination of oxide ions is advantageous in reducing localization behaviour of $2p$ electrons on oxide ions, because there is no nonbonding level on an oxygen in the coordination structure. This may explain the fact that Cu_2O is the known p-type oxide, albeit having strong colour^{8,9}.

Another structural requirement relates to the bandgap. Direct interaction between d^{10} electrons on neighbouring Cu^+ ions may reduce the bandgap. As we are currently interested in transparent materials, an approach is needed to enlarge the bandgap so that there is no absorption in the visible. This is effectively done by lowering the dimension of crosslinking of Cu^+ ions.

Taking into consideration the above requirements, we have selected as a candidate material an aluminate of Cu^+ , CuAlO_2 , with the delafossite structure^{10–13}, which is shown in Fig. 1. Local symmetries around Cu^+ and O^{2-} ions in this phase are the same as those in Cu_2O , except that the nearest-neighbouring cations of the oxide ion are one Cu^+ and three Al^{3+} in the delafossite. The structure

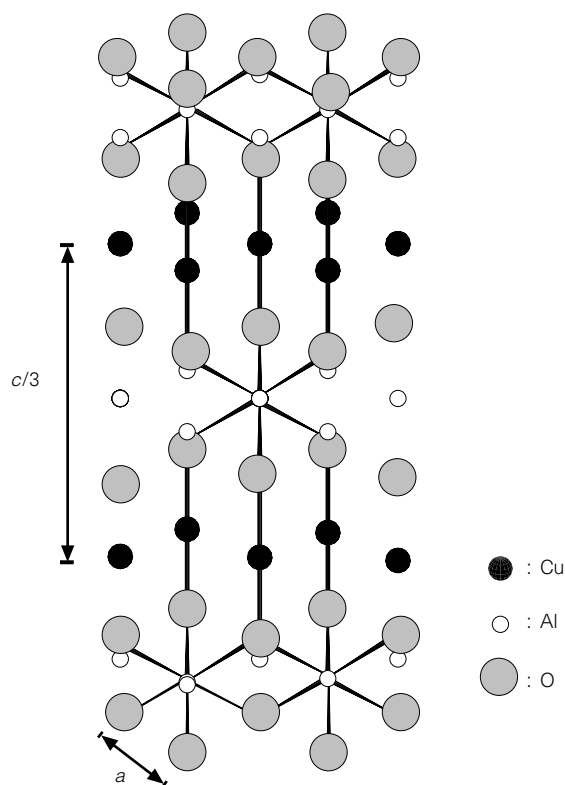


Figure 1 Crystal structure of CuAlO_2 delafossite consisting of stacked layers of $-\text{O}-\text{Al}-\text{O}-\text{Cu}-\text{O}-$ in this order along the c -axis.

can be viewed as a layer structure with the sequence $\text{Cu}-\text{O}-\text{Al}-\text{O}-\text{Cu}$ along the c -axis, and the three-dimensional crosslinking of the Cu^+ network is reduced to two-dimensional. This may give rise to a larger bandgap for the delafossite than for Cu_2O . Benko and Koffyberg have reported a positive sign of Seebeck coefficient for a sintered body of the material¹⁴. However, because of the low conductivity ($1.7 \times 10^{-3} \text{ S cm}^{-1}$) and mobility ($\leq 1 \times 10^{-1} \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$), they proposed a hopping mechanism for hole transport.

Polycrystalline thin films of the delafossite were prepared by laser ablation. Sintered disks were used as a target material. The disks

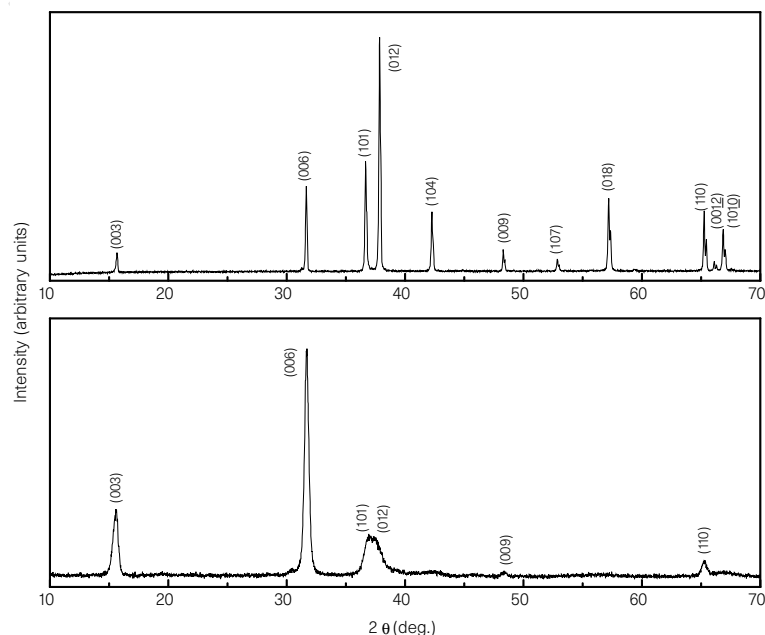


Figure 2 X-ray powder diffraction patterns of CuAlO_2 delafossites. The upper trace is for powders of the sintered disk, and the lower trace is for the thin film prepared by laser ablation. All the diffraction peaks observed could be indexed by assuming the delafossite structure.

were prepared by using solid-state reactions of Cu_2O and Al_2O_3 (ref. 12). Preparation conditions used in the laser ablation were as follows: excitation was by pulses from a KrF excimer laser, 248 nm, 20 Hz, and 200 mJ cm^{-2} (Lambda Physik Compex 102); substrate was a single-crystalline sapphire with c surface orientation (0001); temperature, 700°C ; base pressure, 1×10^{-9} torr; atmosphere, oxygen with a pressure of 0.1 torr. Film thickness was estimated by using a stylus (Sloan Dektak 3030).

The thin films were subjected to the following investigations: chemical analysis by inductively coupled plasma emission spectroscopy; field-emission scanning electron microscope (FESEM) observation of texture; X-ray powder diffraction; optical absorption measurements in the ultraviolet, visible and near-infrared; and measurements of d.c. conductivity in a temperature range of 300–77 K, and of Seebeck coefficients and Hall voltages at around room temperature.

Chemical analyses of CuAlO_2 delafossite samples by the inductively coupled plasma method gave the Cu/Al atomic ratio in the sintered body and thin film as 1.0 and 1.0 respectively. These were in satisfactory agreement with the stoichiometric ratio. Estimation of nonstoichiometry due to defects was not possible by this method. Film thickness was around 500 nm.

Figure 2 shows the X-ray powder diffraction patterns of the sintered body and as-deposited thin film. All diffraction peaks seen in the patterns of both materials could be indexed by assuming the delafossite structure, belonging to the $R3m$ space group. Comparison of the patterns shows that the thin film tended to be oriented on the (001) surface. Although not displayed, FESEM photographs of the surface and cross-section of the thin film showed that the grains were small, around 100 nm in diameter. This might be the origin of the broadened diffraction peaks of the thin film.

The optical absorption spectrum of the thin film in the visible and near-infrared region is displayed in Fig. 3. The fundamental absorption starts from about 350 nm. The inset shows a plot of $(\alpha h\nu)^2$ against $h\nu$, from which the direct allowed bandgap can be estimated; it was found to be about 3.5 eV. However, a diffuse reflectance spectrum showed a small absorption band around 700 nm, and the indirect bandgap may be smaller than 3.5 eV (ref. 14).

The temperature dependence of electrical conductivity is shown in Fig. 4. It is immediately evident that there is a semiconductive temperature dependence. The roughly estimated activation energy for the higher-temperature region was 0.2 eV, which is far smaller than half of the bandgap. This might be an indication of hole transport in the valence band, the hole being thermally activated from an acceptor. At lower temperatures, below ~ 150 K, conductivities followed almost a $T^{-1/4}$ rule, as seen in the inset. The

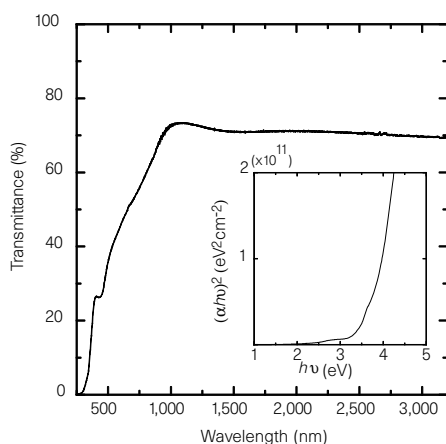


Figure 3 Transmittance spectra of CuAlO_2 thin film. The film thickness was about 500 nm. Inset shows plot of $(\alpha h\nu)^2$ against $h\nu$ for estimation of direct allowed optical gap. The estimated gap was 3.5 eV.

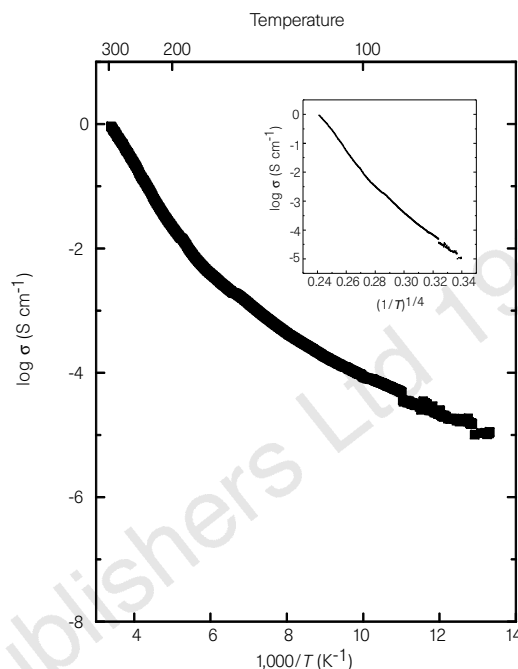


Figure 4 Temperature dependence of electrical conductivity for the CuAlO_2 thin film. The activation energy estimated in the higher-temperature region was about 0.2 eV. Inset shows $T^{-1/4}$ plot.

temperature dependence suggests the presence of gap states close to the Fermi level, and that a variable-range hopping mechanism is dominant at the lower temperatures. The conductivity at room temperature was $0.95 \times 10^{-1} \text{ S cm}^{-1}$. The Hall coefficient, measured by van der Pauw electrode configuration, gave $+48.6 \text{ cm}^3 \text{ C}^{-1}$, suggesting the conduction to be p-type. Combining the Hall coefficient and conductivity measurements resulted in a carrier density of $1.3 \times 10^{17} \text{ cm}^{-3}$ and Hall mobility of the positive holes of $10.4 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$. The Seebeck coefficient was found to be $+183 \mu\text{V K}^{-1}$, again suggesting p-type conductivity.

No intentional doping of positive holes was carried out in this work. Therefore, the origin of the positive holes is not clear at present. The most probable explanation might be excess oxygen. There are two possible models for this excess oxygen: one is cation vacancy and another is an interstitial oxygen. In the case of Cu_2O , the origin of the positive holes was reported to be the ionized Cu vacancies and ionized interstitial oxygens¹⁵. These might be the origin of the positive holes in the delafossite. Intentional doping and elucidation of the origin of the positive holes are topics for further research.

The highest conductivity observed in this study was about 1 S cm^{-1} , which was 3–4 orders of magnitude smaller than that of n-type conducting indium–tin oxide. It is therefore obvious that the thin-film fabrication process must be refined to obtain higher conductivity. Substitutional doping might be the most promising way for increasing the conductivity. The observed Hall mobility of $10 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ is comparable to that of ITO and said to be high as a polycrystalline thin film of an ionic crystal with wide bandgap. The hole concentration of the thin film could be increased to the order of 10^{20} cm^{-3} from the observed value of $1.3 \times 10^{17} \text{ cm}^{-3}$ by substitutional doping. This would probably give rise to two orders of magnitude higher conductivity, even after allowing for the fact that the increase in the hole concentration may lower the Hall mobility to some extent. □

Received 16 June; accepted 9 September 1997.

1. Hamberg, I. & Granqvist, C. G. Evaporated Sn-doped In_2O_3 films: basic optical properties and applications to energy-efficient windows. *J. Appl. Phys.* **60**, R123–R159 (1986).

2. Chopra, K. L., Major, S. & Pandya, D. K. Transparent conductors—A status review. *Thin Solid Films* **102**, 1–46 (1983).
3. Ueda, N. *et al.* New oxide phase with wide band gap and high electroconductivity, MgIn_2O_4 . *Appl. Phys. Lett.* **61**, 1954–1955 (1992).
4. Kawazoe, H. *et al.* Generation of electron carriers in insulating thin film of MgIn_2O_4 spinel by Li^+ implantation. *J. Appl. Phys.* **76**, 7935–7941 (1994).
5. Cava, R. J. *et al.* GaInO_3 : A new transparent conducting oxide. *Appl. Phys. Lett.* **64**, 2071–2072 (1994).
6. Seager, C. H., McIntyre, D. C., Warren, W. L. & Tuttle, B. A. Charge trapping and device behavior in ferroelectric memories. *Appl. Phys. Lett.* **68**, 2660–2662 (1996).
7. Prins, M. W. J. *et al.* A ferroelectric transparent thin-film transistor. *Appl. Phys. Lett.* **68**, 3650–3652 (1996).
8. Hayashi, M. & Katsuki, K. Absorption spectrum of cuprous oxide. *J. Phys. Soc. Jpn* **5**, 380–381 (1950).
9. Kleinman, L. & Mednick, K. Self-consistent energy band of Cu_2O . *Phys. Rev. B* **21**, 1549–1553 (1980).
10. Hahn, H. & Lorent, C. Versuche zur Darstellung ternärer Oxyde des Aluminiums, Galliums und Indiums mit einwertigem Kupfer und Silber. *Z. Anorg. Allg. Chem.* **279**, 281–288 (1955).
11. Shannon, R. D., Rogers, D. B. & Prewitt, C. T. Chemistry of noble metal oxide. 1. Syntheses and properties of ABO_2 delafossite compounds. *Inorg. Chem.* **10**, 713–718 (1971).
12. Ishiguro, T., Kitazawa, A., Mizutani, N. & Kato, M. Single-crystal growth and crystal structure refinement of CuAlO_2 . *J. Solid State Chem.* **40**, 170–174 (1981).
13. Ishiguro, T., Ishizawa, N., Mizutani, N. & Kato, M. Charge-density distribution in crystals of CuAlO_2 with $d-s$ hybridization. *Acta Cryst.* **B39**, 564–569 (1983).
14. Benko, F. A. & Koffyberg, F. P. Opto-Electronic Properties of CuAlO_2 . *J. Phys. Chem. Solids* **45**, 57–59 (1984).
15. Porat, O. & Riess, I. Defect chemistry of Cu_2O at elevated temperatures. Part 2. Electrical conductivity, thermoelectric power and charged point defects. *Solid State Ionics* **81**, 29–41 (1995).

Correspondence should be addressed to H.K. (e-mail: h.kawazoe@rlem.titech.ac.jp).

Low-field magnetoresistance in the pyrochlore $\text{Ti}_2\text{Mn}_2\text{O}_7$

H. Y. Hwang & S.-W. Cheong

Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey 07974, USA

The discovery of colossal magnetoresistance in perovskite manganites derived from LaMnO_3 has renewed interest in these and related materials for possible technological applications^{1–7}. But the high magnetic fields and narrow temperature windows where the large magnetoresistive responses are observed present daunting limitations. A possible avenue for overcoming these limitations is to improve the low-field response by modifying the manganites to enhance the contribution to magnetoresistance from spin-polarized tunnelling of the conduction electrons; such tunnelling, which can be very sensitive to an applied magnetic field, takes place across grain boundaries (natural or artificial) or between the manganite planes of the layered derivatives of the perovskite compounds^{8–13}. Here we show that low-field ($H < 1 \text{ kOe}$) grain-boundary magnetoresistance can also be realized in a non-perovskite magnetoresistive oxide, the pyrochlore $\text{Ti}_2\text{Mn}_2\text{O}_7$. Moreover, this low-field response, which persists for all temperatures below the transition to the ferromagnetic state, does not show the strong temperature-dependent decay characteristic of the perovskite-based systems. We suggest that this improved response is in part due to weaker electron-spin coupling and the lack of strong electron–lattice interactions in the pyrochlore.

We studied polycrystalline samples of $\text{Ti}_2\text{Mn}_2\text{O}_7$ which were made by heating at $\sim 850^\circ\text{C}$ and $\sim 35 \text{ kbar}$ for 1.5 h in a piston-cylinder high-pressure apparatus⁷. Figure 1 shows the temperature-dependent resistivity $\rho(T)$ of $\text{Ti}_2\text{Mn}_2\text{O}_7$ in an applied magnetic field of 5 T and with no applied field. The magnetoresistance (MR) is also shown in Fig. 1; it was measured in the longitudinal geometry (magnetic field $\mathbf{H}$ parallel to current $\mathbf{J}$) with the standard four-probe technique. The abrupt decrease in ρ at 120 K coincides with the onset of ferromagnetism. Near the Curie temperature T_C , colossal magnetoresistance (CMR) is observed, but we focus here on the residual MR below T_C in the metallic ferromagnetic state. In Fig. 2 we present a comparison of the MR below T_C (top panel) and the field-dependent magnetization M (bottom panel), measured with a SQUID magnetometer. Considering first the MR; the most striking feature is the large low-field drop in $\rho(H)$, which is apparent for all $T < T_C$. With increasing temperature, the relative size of this low-field feature is reduced, and the bulk intrinsic MR increases,

becoming dominant near T_C . The lower panel of Fig. 2 shows $M(H)$ for a few representative temperatures. The behaviour is typical of a magnetically soft ferromagnet, with a quick rise at low fields as the magnetic domains are aligned with the applied field, and then saturating above 1,000 Oe. The field scale for the low-field MR is the same as that for magnetic-domain rotation. In samples with different geometric demagnetization factors, these features shift together, and both transverse and longitudinal MR are similar.

We attribute low-field MR to the grain boundaries (as opposed to scattering by the Bloch magnetic-domain walls in the bulk), on the basis of a study of samples of widely differing grain sizes. Owing to slight variations in the growth conditions, samples can be obtained with grain sizes ranging from the micrometre scale (the data discussed above), to samples composed of crystallites which are visible to the naked eye. The size of the low-field MR at 5 K ranges from $\Delta\rho/\rho_0 \approx 30\%$ for the smallest-grain sample (where ρ_0 is the zero-field resistivity) to $\sim 1\%$ in the largest-grain samples, although the magnetization is identical in all samples (Fig. 2). This is consistent with the association of the low-field MR with the grain boundaries, not with the intrinsic Bloch walls in the bulk material. Also, one would expect history-dependent MR if it was associated with the movement of Bloch walls in the bulk, but our measurements are completely reproducible, being independent of field cooling, zero-field cooling and so on.

Extrapolating the low-field MR to $H = 0$ gives MR^* ; the temperature dependence of this latter quantity is given in Fig. 3. The error bars reflect the uncertainty in extrapolating the high-field MR to zero field, the subtraction of which gives MR^* . For comparison, we have also plotted various powers of M taken at 1,000 Oe, and it appears that M^3 most closely approximates the data. The importance of this temperature dependence becomes clear when our data are compared to results obtained from the perovskite manganites. Also shown in Fig. 3 is the temperature dependence of MR^* in polycrystalline $\text{La}_{2/3}\text{Sr}_{1/3}\text{MnO}_3$ (LSMO)⁹, as well as the temperature dependence of the tunnelling magnetoresistance (TMR) observed in a thin-film tri-layer tunnel junction with SrTiO_3 (STO) as the insulating layer¹⁰. Despite the difference between these two systems, they share a very similar temperature dependence, characterized by a rapid decrease in magnitude far below T_C . By contrast, a much weaker temperature dependence of the low-field MR is demonstrated in $\text{Ti}_2\text{Mn}_2\text{O}_7$.

It appears that the low-field MR can be viewed as analogous to the TMR observed in the perovskite manganites. There are, however, a number of differences that affect the understanding of this phenomenon in the pyrochlore. In perovskite manganites, the bulk

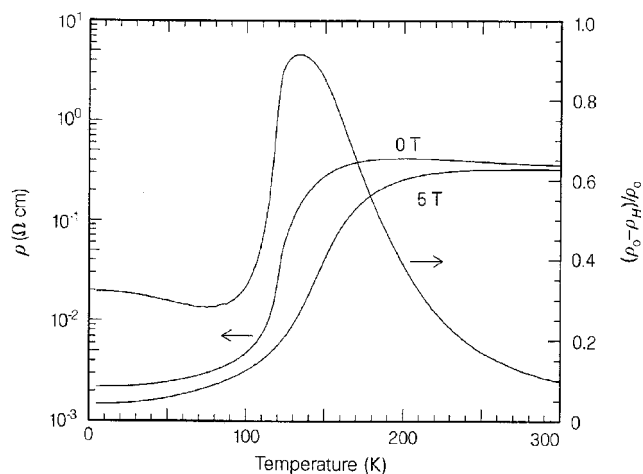


Figure 1 The temperature-dependent resistivity $\rho(T)$ of $\text{Ti}_2\text{Mn}_2\text{O}_7$ in 0 and 5 T applied fields. Also shown is the MR, $(\rho_0 - \rho_H)/\rho_0$. We focus on the low-temperature MR below the ferromagnetic transition at 120 K.

magnetotransport properties can be understood within the double-exchange framework^{14,15}. Ferromagnetic coupling between Mn^{3+} and Mn^{4+} arises from the hopping of an electron between the two partially filled d shells with strong intra-atomic Hund's coupling J_H . Because the alignment of spins on neighbouring sites favours electronic motion, the ferromagnetic ground state maximizes the electron itineracy energy. In the pyrochlore, however, the magnetism does not arise from this type of double exchange because all of the Mn ions are in the Mn^{4+} state¹⁶. Instead the ferromagnetism is predominantly mediated by superexchange, and the conduction electrons are probably derived from the Tl 6s band¹⁷. Thus in $\text{Tl}_2\text{Mn}_2\text{O}_7$ there are two separate electronic systems: a magnetic sublattice of Mn–O and a conducting sublattice of Tl–O indirectly coupled to it.

Another important difference between the two systems is the role of strong electron–lattice coupling in the perovskites. The Mn^{3+} ions tend to distort the local oxygen positions so as to minimize the energy of the lone electron in the doubly degenerate e_g level (the Jahn–Teller effect). This competition between localizing lattice distortions and electron itineracy has been shown experimentally and theoretically to be a key feature in enhancing MR in the

perovskites^{18–20}. In contrast to this is the absence of strong electron–phonon coupling in the pyrochlore. As all the Mn are tetravalent, there are no Jahn–Teller energy gains via lattice distortions. Consistent with this is the lack of a volume anomaly at T_C as determined by synchrotron X-ray studies on the pyrochlore (D. E. Cox *et al.*, personal communication). In the perovskites, the large volume contraction at T_C is one of many signatures of strong electron–lattice coupling^{19,20}.

We now consider transport across a single grain boundary (or an artificial boundary in a thin-film trilayer), shown schematically in Fig. 4. In zero applied field, the magnetization M on either side of the barrier tends to be antialigned by dipole coupling, and thus a majority spin conduction electron tends not to cross the barrier because it becomes a minority carrier on the other side. Once the two sides are aligned by an external magnetic field, the majority spin conduction electrons can cross the barrier and propagate on the other side. In the perovskites, at low temperatures, the conduction electrons are highly spin-polarized because the full magnetic moment is ordered and J_H is greater than b , the electronic bandwidth. This unusual situation arises despite the high carrier concentration (nominally ~ 0.3 carriers per formula unit). In the

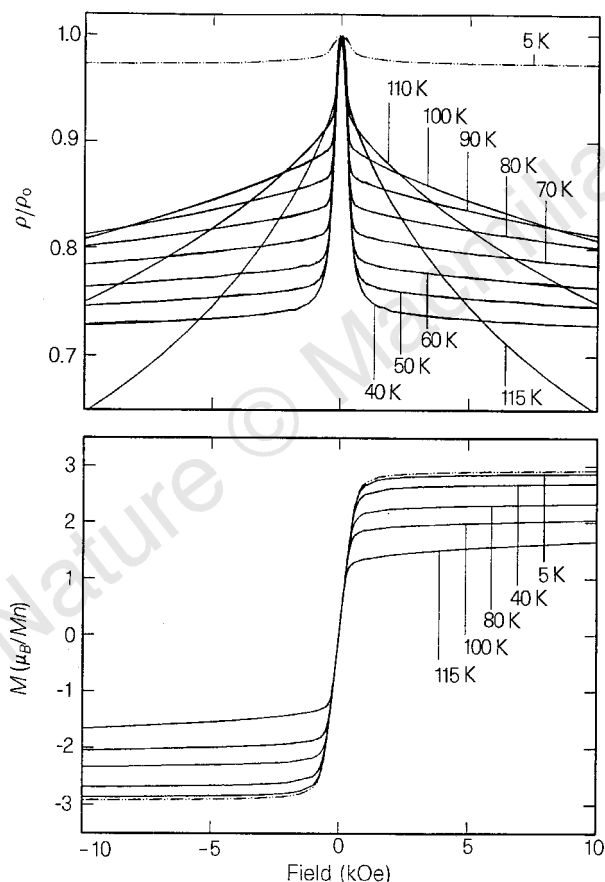


Figure 2 The top panel displays the field-dependent resistivity normalized to the zero field value $\rho(H)/\rho_0$, for a number of temperatures below $T_C = 120$ K. Solid lines are for the smallest grain sample with maximal MR (shown in Fig. 1), and broken lines are for the largest grain sample with minimal MR. Two regions are delineated: a sharp low-field drop associated with magnetic domain rotation, and high-field residual MR. As T_C is approached, the CMR behaviour begins to dominate. The bottom panel gives the field-dependent magnetization for a few representative temperatures. The similarity between the two samples at 5 K demonstrates their identical bulk magnetic natures. For both panels, the data are shown for magnetic field sweeps in both directions.

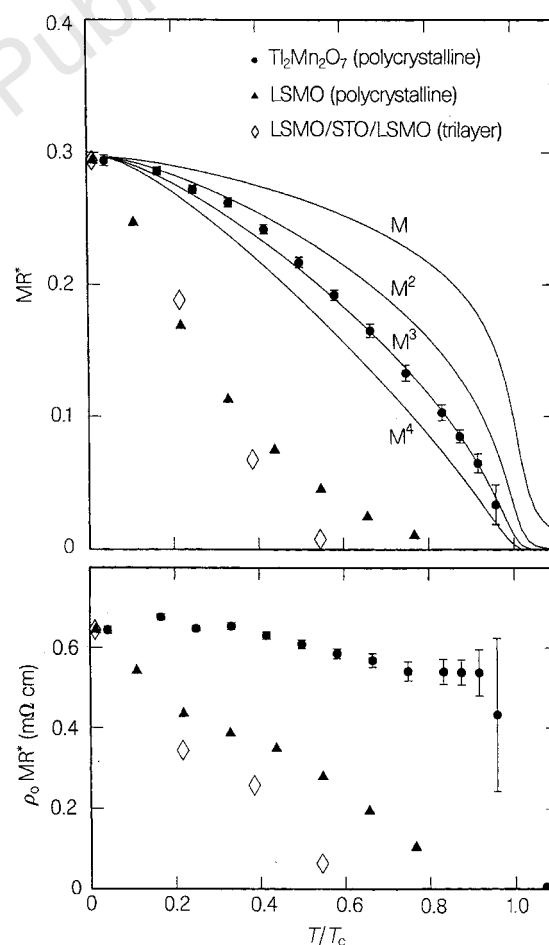


Figure 3 The top panel displays the temperature dependence (normalized to T_C) of the magnitude of the low-field MR extrapolated to $H = 0$, $\text{MR}^*(T)$, for three samples: polycrystalline $\text{Tl}_2\text{Mn}_2\text{O}_7$ (this work), polycrystalline $\text{La}_{2/3}\text{Sr}_{1/3}\text{MnO}_3$ (LSMO) (from ref. 9), and a thin-film trilayer tunnel junction of $\text{La}_{2/3}\text{Sr}_{1/3}\text{MnO}_3$ with SrTiO_3 as the tunnelling barrier (from ref. 10). The perovskite data have been normalized to the zero-temperature magnitude of MR^* for $\text{Tl}_2\text{Mn}_2\text{O}_7$. Also shown (normalized) are various powers of the magnetization M for $\text{Tl}_2\text{Mn}_2\text{O}_7$ at 1,000 Oe, just above the field scale for magnetic domain rotation. The bottom panel gives the absolute magnitude of MR^* , that is $\rho_0 \text{MR}^*$, for the three samples. The perovskite data have again been normalized to the value for $\text{Tl}_2\text{Mn}_2\text{O}_7$.

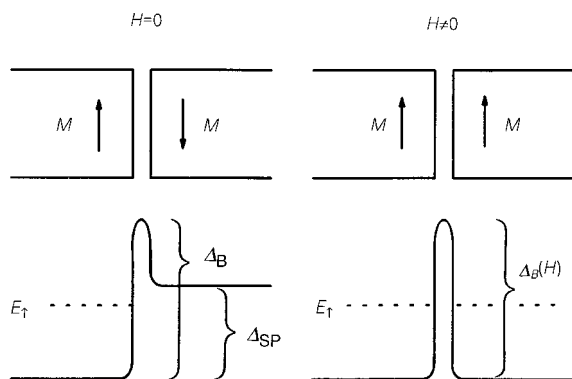


Figure 4 Schematic energy-level diagram associated with transport across a grain boundary. On the left, in zero applied magnetic field, the grains tend to be magnetically antialigned by dipole coupling. (This simplified representation neglects the effects of magnetocrystalline anisotropy and geometric restrictions on dipole coupling in three dimensions.) There is a potential barrier at the grain boundary Δ_B . The energy level is given for a majority carrier of the left grain, which corresponds to a minority carrier in the right grain. (The potential has been depicted for full spin polarization, $\Delta_{SP} > E_T$.) On the right, in a vertical applied magnetic field, the grains are magnetically aligned. Thus the spin orientation of the majority carrier is the same in both grains, with a potential barrier between them.

pyrochlore, a large degree of spin-polarization is primarily due to the extremely small number of carriers (estimated at 0.006×10^{21} carriers per formula unit from Hall effect measurements⁶). Thus even moderate exchange splitting of Tl 6s-derived bands can fully polarize the conducting electrons.

What, then, can account for the different temperature dependences of the transport process at the grain boundary between the two systems? In the perovskite manganites, the surface magnetism will in general be quite different from that of the bulk, owing to the lower effective magnetic coupling, the tendency of spin canting (an important feature in the perovskite manganites²¹), as well as the disorder due to the surface termination of the crystal structure. Also important is the influence of the Jahn–Teller lattice distortions at the surface. One possible consequence of these effects would be a lower surface T_C and surface spin polarization with respect to the bulk. The hopping electron is strongly coupled to these surface spins by the large on-site J_H , and thus the TMR is susceptible to fast degradation with increasing temperature via the magnetically altered surface spins. In the pyrochlore, however, there is no strong J_H coupling the conducting electrons to the Mn spins, and our results imply less sensitivity to the effects of the surface.

Another distinction can be visualized by considering the temperature dependence of MR^* in absolute units ($\rho_0 MR^*$ shown in the lower panel of Fig. 3). In the perovskite, the absolute magnitude of MR^* diminishes rapidly with increasing temperature, much below T_C . In contrast, the same data for the pyrochlore is virtually constant for all temperatures up to T_C . This difference may indicate that for $Tl_2Mn_2O_7$, the potential barrier Δ_B is less than the electron energy E_T (referring to Fig. 4). In this case the grain boundary transport would not be considered tunnelling, but rather a scattering process across a disordered region acting as a series resistance. This implies that the low-field MR can be made even larger by raising the relative resistance of the grain boundaries with respect to the bulk resistivity, much as a thin-film tunnel junction increases this ratio with respect to bulk polycrystalline samples in perovskite manganites. □

Received 3 April; accepted 26 August 1997.

1. Jonker, G. H. & Van Santen, J. H. Ferromagnetic compounds of manganese with perovskite structure. *Physica* **16**, 337–349 (1950).
2. Searle, C. W. & Wang, S. T. Studies of the ionic ferromagnet $(LaPb)MnO_3$. V. Electric transport and

ferromagnetic properties. *Can. J. Phys.* **48**, 2023–2031 (1970).

3. von Helmholt, R., Wecker, J., Holzapfel, B., Schultz, L. & Samwer, K. Giant negative magnetoresistance in perovskitelike $La_{2/3}Ba_{1/3}MnO_3$ ferromagnetic films. *Phys. Rev. Lett.* **71**, 2331–2333 (1993).
4. Chahara, K., Ohno, T., Kasai, M. & Kozono, Y. Magnetoresistance in magnetic manganese oxide with intrinsic antiferromagnetic spin structure. *Appl. Phys. Lett.* **63**, 1990–1992 (1993).
5. Jin, S. *et al.* Thousandfold change in resistivity in magnetoresistive La–Ca–Mn–O films. *Science* **264**, 413–415 (1994).
6. Shimakawa, Y., Kubo, Y. & Manako, T. Giant magnetoresistance in $Tl_2Mn_2O_7$ with the pyrochlore structure. *Nature* **379**, 53–55 (1996).
7. Cheong, S.-W., Hwang, H. Y., Batlogg, B. & Rupp, L. W. Jr Giant magnetoresistance in pyrochlore $Tl_2Mn_2O_7$. *Solid State Commun.* **98**, 163–166 (1996).
8. Gubkin, M. K., Perekalina, T. M., Bykov, A. V. & Chubarenko, V. A. Investigation of the electrical conductivity of some manganites with the perovskite structure. *Phys. Solid State* **35**, 728–730 (1993).
9. Hwang, H. Y., Cheong, S.-W., Ong, N. P. & Batlogg, B. Spin-polarized intergrain tunneling in $La_{2/3}Sr_{1/3}MnO_3$. *Phys. Rev. Lett.* **77**, 2041–2044 (1996).
10. Sun, J. Z. *et al.* Observation of large low-field magnetoresistance in trilayer perpendicular transport devices made using doped manganite perovskites. *Appl. Phys. Lett.* **69**, 3266–3268 (1996).
11. Kimura, T. *et al.* Interplane tunneling magnetoresistance in a layered manganite crystal. *Science* **274**, 1698–1701 (1996).
12. Gupta, A. *et al.* Grain-boundary effects on the magnetoresistance properties of perovskite manganite films. *Phys. Rev. B* **54**, 15629–15632 (1996).
13. Mathur, N. D. *et al.* Large low-field magnetoresistance in $La_{0.7}Ca_{0.3}MnO_3$ induced by artificial grain boundaries. *Nature* **387**, 266–268 (1997).
14. Zener, C. Interaction between the *d*-shells in the transition metals. II. Ferromagnetic compounds of manganese with perovskite structure. *Phys. Rev.* **82**, 403–405 (1951).
15. Anderson, P. W. & Hasegawa, H. Considerations on double exchange. *Phys. Rev.* **100**, 675–681 (1955).
16. Subramanian, M. A. *et al.* Colossal magnetoresistance without Mn^{3+}/Mn^{4+} double exchange in the stoichiometric pyrochlore $Tl_2Mn_2O_7$. *Science* **273**, 81–84 (1996).
17. Singh, D. J. Magneto-electronic effects in pyrochlore $Tl_2Mn_2O_7$: role of Tl–O covalency. *Phys. Rev. B* **55**, 313–316 (1997).
18. Millis, A. J., Littlewood, P. B. & Shraiman, B. I. Double exchange alone does not explain the resistivity of $La_{1-x}Sr_xMnO_3$. *Phys. Rev. Lett.* **74**, 5144–5147 (1995).
19. Dai, P. *et al.* Experimental evidence for the dynamic Jahn–Teller effect in $La_{0.65}Ca_{0.35}MnO_3$. *Phys. Rev. B* **54**, 3694–3697 (1996).
20. Radaelli, P. G., Marezio, M., Hwang, H. Y., Cheong, S.-W. & Batlogg, B. Charge localization by static and dynamic distortions of the MnO_6 octahedra in perovskite manganites. *Phys. Rev. B* **54**, 8992–8995 (1996).
21. de Gennes, P.-G. Effects of double exchange in magnetic crystals. *Phys. Rev.* **118**, 141–154 (1960).

Acknowledgements. We thank B. Batlogg, Y. B. Kim, P. B. Littlewood, P. Majumdar, A. P. Ramirez and B. I. Shraiman for helpful discussions.

Correspondence and requests for materials should be addressed to H.Y.H.

A combinatorial approach to the discovery and optimization of luminescent materials

Earl Danielson, Josh H. Golden, Eric W. McFarland, Casper M. Reeves, W. Henry Weinberg & Xin Di Wu

Symyx Technologies, 3100 Central Expressway, Santa Clara, California 95051, USA

Combinatorial synthesis and screening of very large numbers of organic compounds has been widely applied in the pharmaceutical industry for drug discovery¹. Recently, combinatorial arrays of inorganic materials with known or potential superconductivity² and giant magnetoresistance³ have been synthesized and screened. The combinatorial approach is particularly well suited to ternary and higher-order inorganic materials, for which efforts to predict basic properties have been unsuccessful⁴. Here we describe an automated combinatorial method for synthesizing and characterizing thin-film libraries of up to 25,000 different materials, on a three-inch-diameter substrate, as candidates for new phosphors. The discovery and development of new compounds for ultraviolet-excited phosphors is of great importance for the development of flat-panel displays⁵ and lighting⁶. As there are no reliable theories to predict the relation between composition and phosphor colour and efficiency, the less than 100 useful commercial phosphor materials have been discovered through one-by-one serial synthesis and testing^{7,8}. Our approach, in contrast, offers rapid screening of many compositions, and it has enabled us to identify a new red phosphor, $Y_{0.845}Al_{0.070}La_{0.060}Eu_{0.025}VO_4$, which has a quantum efficiency comparable or superior to those of existing commercial red phosphors.

Our application of solid state combinatorial chemistry, exemplified by the phosphor materials discussed here, involves the following steps: deposition of multiple solid state materials as thin films in varying ratios on solid substrates in a pre-designed spatial arrangement called a combinatorial library; thermal processing of the library to mix the materials and form either stable or metastable compounds; high-throughput screening of the individual constituents of the library to identify possible 'lead' candidates; design, synthesis, and screening of new focus libraries based upon the compositions of the initial 'leads'; and, finally, synthesis of the optimized 'lead' compounds as powders by conventional synthetic routes, followed by detailed analytical characterization. To implement this methodology we have developed technology that has made possible the high-speed, automated, combinatorial synthesis and screening of unprecedented numbers of different solid state materials. The methodology can be applied to a variety of problems in materials science to increase by several orders of magnitude the rate of new material discovery and optimization.

Typical phosphors are inorganic powders consisting of a polycrystalline host doped with rare earth or transition metal ions, or both. The dopants serve as the luminescent centre (activator) and/or as a sensitizer which absorbs and transfers incident ultraviolet excitation energy to the luminescent centre^{9,10}. We designed a phosphor library to include combinations of host lattice cations from groups IIA and IIIA, with VO_m^{n-} ($m = 3, 4$; $n = 1, 3$) serving as the counter anion, and from the group IIIB and IVB elements (Fig. 1a). In this library, representative compounds are included of large, nearly continuous regions of composition space for the selected hosts and the activator ions, namely Eu, Tb, Tm and Ce.

Combinations containing VO_m^{n-} and $\text{VO}_m^{n-}/\text{Al}$ were added because Al substitution has been shown empirically to increase the luminosity in $\text{Y}_{0.9}\text{Al}_{0.05}\text{Eu}_{0.05}\text{VO}_4$ (ref. 11). A powerful feature of the combinatorial approach is the ease with which such 'educated guesses' can be investigated and extended to analogous chemical systems by including diversity on single libraries.

Libraries were deposited as thin films by electron beam (8 kV) evaporation in vacuum (base pressure $\sim 5 \times 10^{-6}$ torr) on unheated three-inch-diameter silicon wafers. The rate was approximately 0.3 nm s^{-1} as monitored with an *in situ* quartz crystal microbalance. An electron gun with multiple graphite crucible target pockets was used with La_2O_3 , Y_2O_3 , MgO , SrCO_3 , SnO_2 , V, Al_2O_3 , Eu_2O_3 , Tb_4O_7 , Tm_2O_3 and CeO_2 solid pellet targets (purities $\geq 99.99\%$, except V which was $\geq 99.70\%$). A stainless steel primary mask consisting of $230 \mu\text{m}$ square elements spaced $420 \mu\text{m}$ apart was attached to the substrate to separate individual library elements. The spatial variation of materials deposited on the library was created using stationary and movable physical masks to control the thickness of specific evaporants in selected regions of the substrate (Fig. 1a). Two mask sets were used, one with a single 19.1-mm-wide rectangular slit and the other with four 4.8-mm-wide rectangular slits. Four columns were deposited first with constant thicknesses of SnO_2 , V, $\text{Al}_2\text{O}_3 + \text{V}$ (15:8 molar ratio), and Al_2O_3 . Four rows with linearly varying thicknesses of La_2O_3 , Y_2O_3 , MgO and SrCO_3 were then layered on the columns to divide the substrate into 16 host lattice subregions. Finally, within each subregion, columns of rare earths Eu_2O_3 , Tb_4O_7 , Tm_2O_3 and CeO_2 were deposited in linearly varying thicknesses. On this library there were approximately 600 different chemical compositions per square centimetre.

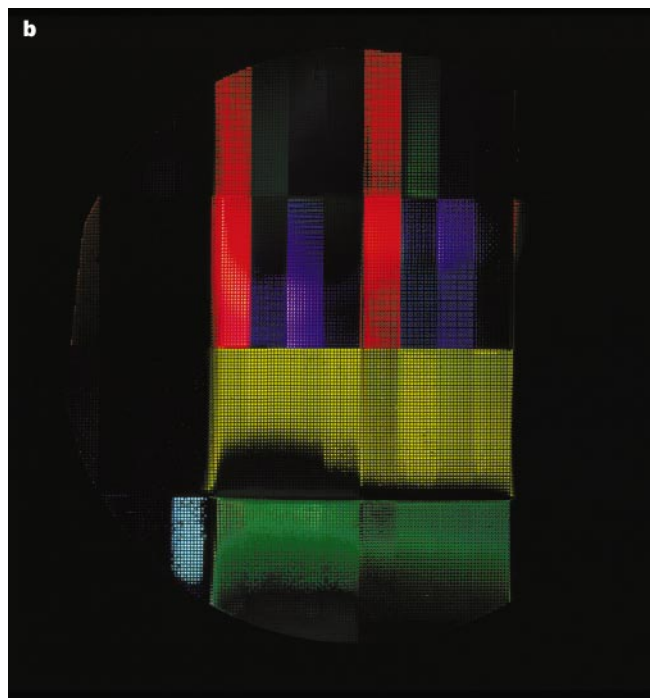
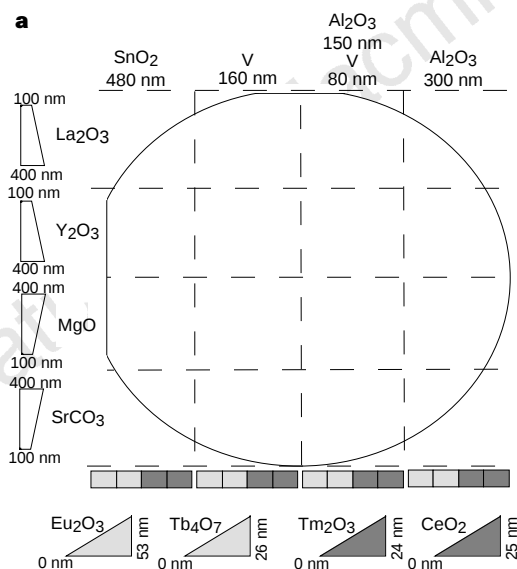
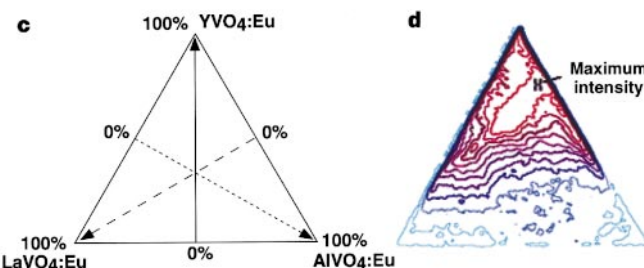


Figure 1 Solid state combinatorial libraries. **a**, Deposition map of diverse phosphor discovery library. Thicknesses of target materials deposited on the 3-inch-diameter substrate are shown. **b**, Photograph of the ultraviolet-excited (254 nm) photoluminescence from the library. **c**, Deposition map of host optimization library exploring all possible stoichiometries of Y, Al and La vanadates formed by varying the host constituents linearly along three axes rotated by 120° . The Eu and V were deposited uniformly over the entire substrate. **d**, False-colour contour plot of the red emission from the ultraviolet-excited (254 nm) photoluminescence of the $\text{Y}_{0.95-m-n}\text{Al}_n\text{La}_m\text{Eu}_{0.05}\text{VO}_4$ library. Contours represent linear changes in relative intensity from 100% (red) to roughly 50% (turquoise).



The constituents of each library element were deposited in layers. Interplanar mixing and formation of the inorganic compounds was accomplished by subsequent oxidative thermal processing at a variety of temperatures (see Methods). Because of the large variation in material composition over diverse libraries, single libraries will typically require a number of different processing conditions to explore the set of chemical phases and ensure adequate mixing of layered depositions. Nonetheless, even non-optimized processing under a limited number of conditions provides considerable new knowledge (see below).

High-throughput screening of combinatorial libraries for ultraviolet-excited photoluminescence was performed by imaging the visible emission of the library with a CCD camera while exciting luminescence with a 254 nm source (see Methods). A colour photograph of the visible emission from the library is shown in Fig. 1b. A quantitative measure of chromaticity was calculated from three images obtained using, respectively, red, green and blue tristimulus emission filters¹². Deep powder samples of red, green and blue commercial phosphors, namely, $\text{Y}_2\text{O}_3\text{:Eu}$ (10%), ZnS:Cu , Al and ZnS:Ag , Cl , were imaged under identical conditions to provide references and calibration for the chromaticity and luminosity of each library member. Preliminary identification of the phosphors that had highest extrinsic efficiency with desirable red,

green and blue chromaticity was performed by automatically searching the imaging data for those regions of the library with the highest luminance within selected ranges of chromaticity coordinates. Table 1 lists representative compositions of selected red, green and blue phosphors together with their relative intensities, CIE chromaticity coordinates¹² and relative spatial locations on the library.

Combinatorial exploration involves several iterations to optimize composition and processing for a given application. The highest-efficiency materials with desirable chromaticity identified in the initial high-density exploration library were red phosphors (Eu^{3+} doped) with $\text{Y}_{1-m}\text{Al}_m\text{VO}_4$ as the host. To investigate and optimize these 'lead' compounds, a second library was designed and synthesized to focus on the best host compositions. This library included La as a third Group III host. Selected compositions of yttrium vanadates doped with lanthanum have been previously described as efficient lamp and cathodoluminescent phosphors¹³. Both Eu_2O_3 and V were deposited uniformly (26.3 nm and 189.6 nm, respectively) over the entire substrate followed by Y_2O_3 , Al_2O_3 and La_2O_3 , deposited as linear gradients along three axes rotated by 120° (Fig. 1c). The comprehensive range of host compositions explored in this library represents all possible $\text{Y}_{0.95-m-n}\text{Al}_n\text{La}_m\text{Eu}_{0.05}\text{VO}_4$ compounds. A contour map of the ultraviolet-excited red photoluminescent

Table 1 High-throughput screening results for extrinsic phosphor properties

CIE Chromaticity (x, y)	Ranks in library within colour group	Relative luminosity	Location in library (i, j)	Elemental components (molar ratio)
Red ($x = 0.6-0.7, y = 0.3-0.4$)				
(0.65, 0.35)	1 (out of 1,754)	1.00	72, 58	Y:V:Eu (0.34:0.60:0.06)
(0.65, 0.35)	4	0.95	73, 58	Y:V:Eu (0.35:0.59:0.06)
(0.64, 0.35)	149	0.53	66, 105	Y:V:Eu (0.37:0.29:0.30:0.04)
(0.64, 0.34)	380	0.27	26, 110	Al:La:V:Eu (0.34:0.31:0.28:0.07)
(0.64, 0.35)	516	0.21	4, 606	La:V:Eu (0.51:0.43:0.06)
Green ($x = 0.0-0.4, y = 0.4-0.8$)				
(0.40, 0.51)	1 (out of 121)	1.00	102, 140	Al:Mg:V:Ce (0.21:0.61:0.17:0.01)
(0.38, 0.53)	1	1.00	103, 115	Al:Mg:V:Tb (0.22:0.60:0.17:0.01)
(0.37, 0.51)	6	0.73	144, 54	Mg:V:Eu (0.34:0.63:0.03)
(0.36, 0.44)	44	0.59	51, 119	Al:Y:V:Tb (0.43:0.19:0.35:0.03)
(0.35, 0.41)	93	0.50	42, 117	Al:La:V:Tb (0.32:0.26:0.40:0.02)
Blue ($x = 0.0-0.3, y = 0.0-0.2$)				
(0.18, 0.09)	1 (out of 728)	1.00	82, 78	Y:V:Tm (0.41:0.57:0.02)
(0.18, 0.08)	3	0.94	84, 78	Y:V:Tm (0.42:0.56:0.02)
(0.19, 0.11)	29	0.76	68, 126	Al:Y:V:Tm (0.38:0.31:0.30:0.01)
(0.19, 0.11)	37	0.73	70, 126	Al:Y:V:Tm (0.37:0.30:0.32:0.01)
(0.21, 0.13)	307	0.32	70, 80	Y:V:Tm (0.35:0.63:0.02)

Representative high-throughput screening results for extrinsic phosphor properties obtained from the combinatorial library of Fig. 1. An ultraviolet-filtered (330 nm, LWP, Andover) CCD camera (TE-1242E, Princeton Instruments) was used to image the library. A 1:2 reduction onto the 25.9 mm $\times$ 27.5 mm active area of the CCD was achieved using a 120 mm lens (Apo-Rodagon-D, Rodenstock). The library was imaged in several sections. Two 254 nm broadband ultraviolet lamps (UVP, Inc. model UVGL-58) were used for excitation ($\sim 1.1 \text{ mW cm}^{-2}$). The intensities from the three images obtained with the red (602 nm), green (553 nm) and blue (461 nm) bandpass tristimulus emission filters (Andover Corp., Salem, NH) were weighted and combined to calculate the (x, y) CIE chromaticity coordinates¹². The known CIE coordinates of the three standard phosphors (red, $\text{Y}_{10}\text{O}_5\text{S:Eu}_{0.1}$, Osram-Sylvania type 1154, $x = 0.629, y = 0.350$; green, ZnS:Cu, Al , General Electric 118-2-16, $x = 0.266, y = 0.576$; blue, ZnS:Ag, Cl , Osram-Sylvania type 1332, $x = 0.15, y = 0.05$) were used to determine the weighting factors¹². The chromaticity data were searched and those compounds with coordinates within specified ranges of CIE coordinates, red ($x = 0.6-0.7, y = 0.3-0.4$), green ($x = 0.0-0.4, y = 0.4-0.8$), and blue ($x = 0.0-0.3, y = 0.0-0.2$), were ranked according to relative extrinsic efficiency by colour group. The observed brightness or extrinsic efficiency must be used with caution unless comparing samples with identical reflectivity and absorption. Although photometric errors in relative efficiencies were less than 4% film morphological variations, light piping within the films and other factors contributed to an overall uncertainty of about $\pm 20\%$. Relative errors in chromaticity are about $\pm 15\%$.

intensity measured from the library is shown in Fig. 1d. A clearly defined region of relatively high extrinsic efficiency is observed in the intensity map. The maximum intensity with red chromaticity suitable for a commercial material corresponded to the phosphor composition $Y_{0.82}Al_{0.07}La_{0.06}Eu_{0.05}VO_4$. The activator concentration in the lead host was subsequently optimized with a third library with variable Eu^{3+} , that is $Y_{0.87-m}Al_{0.07}La_{0.06}Eu_mVO_4$. Spatial variations of Eu^{3+} concentration between 0 and 20% were produced with a single movable mask. Screening of this optimization library identified 2.5% Eu as the most efficient dopant composition (Fig. 2). To our knowledge $Y_{0.845}Al_{0.07}La_{0.06}Eu_{0.025}VO_4$ represents a new phosphor composition with improved red chromaticity ($x = 0.67$, $y = 0.32$)

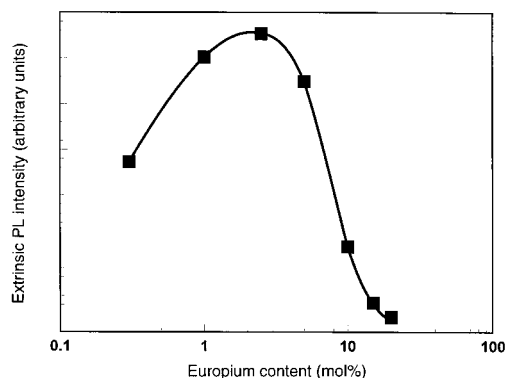


Figure 2 Extrinsic photoluminescence (PL) efficiency measured from an optimization library of $Y_{0.87-m}Al_{0.07}La_{0.06}Eu_mVO_4$ to determine the optimal Eu dopant concentration. The curve is meant as a guide to the eye.

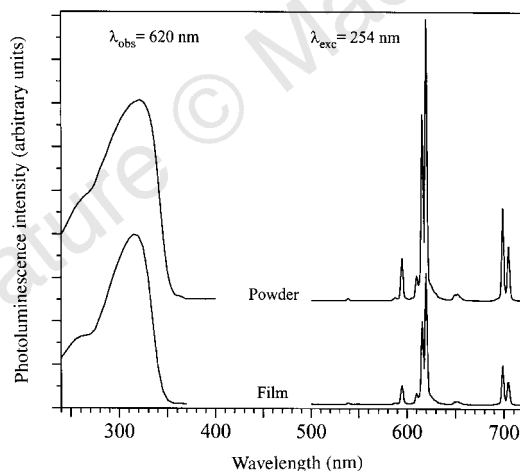


Figure 3 Excitation (left) and emission spectra (right) obtained from the $Y_{0.845}Al_{0.07}La_{0.06}Eu_{0.025}VO_4$ phosphor. Spectroscopic features of the powder (top) are identical to those from the thin-film library (bottom). Spectroscopy and quantum efficiency measurements were performed with a Spex-222, Fluorolog (Hg/Xe) lamp. All measurements were performed at room temperature with a spectral resolution of 0.9 nm. Identical source and detector solid angles were used. The absolute source intensity, $I_0(\lambda) = S(\lambda)/(PMT(\lambda)\theta_{NaSal})$, was determined from the measured photomultiplier signal, $S(\lambda)$, corrected for the efficiency, $PMT(\lambda)$, the fractional solid angle, θ , and the quantum efficiency of the sodium salicylate (NaSal) quantum counter, QE_{NaSal} , which is nearly constant between 250 and 400 nm. The detector response was calibrated with a Labsphere certified reflectance standard, $D(\lambda) = R(\lambda)$. The efficiency, Q , at a specific excitation wavelength was then determined as $Q_{\lambda_{exc}} = I_{em}/I_0(\lambda_{exc})$. Reflectance corrections were applied ratiometrically from the known intrinsic efficiency of a $Y_{1.95}Eu_{0.05}O_3$ standard, $QE_{std}(254\text{ nm}) = 0.91$, and measured reflectances at the same excitation wavelength, R_{λ} , that is, $QE_{sample,\lambda_{exc}} = QE_{std}(Q_{sample,\lambda_{exc}}/Q_{std,\lambda_{exc}})(1 - R_{std,\lambda_{exc}})/(1 - R_{sample,\lambda_{exc}})$.

compared with the more orange standard commercial red phosphor, $Y_{1.95}O_3Eu_{0.05}$ ($x = 0.64$, $y = 0.35$).

After the lead compound was identified and the host lattice and activator concentration were optimized on combinatorial libraries, the resulting $Y_{0.845}Al_{0.07}La_{0.06}Eu_{0.025}VO_4$ and other similar compositions were prepared as powders by conventional methods (see Methods). X-ray diffraction of the powder sample verified that the $Y_{0.845}Al_{0.07}La_{0.06}Eu_{0.025}VO_4$ consisted of a single-phase material isostructural with YVO_4 (JCPDS card No. 17-0341). The spectrum showed the expected small shifts in the lattice parameters consistent with substitutions of La and Al for Y. High-resolution excitation and emission spectroscopy was performed on both the powder and a 1 mm² region of the $Y_{0.87-m}Al_{0.07}La_{0.06}Eu_mVO_4$ thin-film library at the location corresponding to 2.5% Eu (see Methods). The spectra showed identical absorption and emission properties (Fig. 3). Similar comparisons were made for several additional thin films and their corresponding powders from other libraries and non-optimal regions of the focus library, including $Y_{0.845}Al_{0.06}La_{0.07}Eu_{0.05}VO_4$, $Y_{0.88}Al_{0.07}Eu_{0.05}VO_4$, $Y_{2-m}O_3Eu_m$ ($m = 0.01, 0.05, 0.10$) and $La_{0.38}Ce_{0.47}Tb_{0.15}PO_4$. In all cases the bulk powder spectra were indistinguishable from the thin films, indicating that the host-dopant interactions were identical. The intrinsic quantum efficiencies of $Y_{0.845}Al_{0.07}La_{0.06}Eu_{0.025}VO_4$ and $Y_{0.82}Al_{0.07}La_{0.06}Eu_{0.05}VO_4$ powder samples were 0.83 ± 0.06 and 0.78 ± 0.05 , respectively, consistent with the trend observed in the extrinsic efficiency shown in Fig. 2 (ref. 14). Furthermore, the new phosphor, $Y_{0.845}Al_{0.07}La_{0.06}Eu_{0.025}VO_4$, discovered with the combinatorial methodology and synthesized in our laboratory in powder form, is of superior red chromaticity and has an efficiency comparable to that of 0.86 ± 0.04 measured from the standard commercial phosphor, $Y_{1.95}Eu_{0.05}O_3$, but with half the concentration of the expensive rare earth. $Y_{0.95}Eu_{0.05}O_4$ is also used commercially because of its improved chromaticity, but the efficiency is relatively low, 0.68 (ref. 15). Our new $Y_{0.845}Al_{0.07}La_{0.06}Eu_{0.025}VO_4$ has the chromaticity advantage of the vanadate with higher efficiency and lower Eu concentration.

Although, as expected, the powder sample of $Y_{0.845}Al_{0.07}La_{0.06}Eu_{0.025}VO_4$ had considerably higher relative extrinsic brightness than the thin film of the same composition, the intrinsic efficiencies were comparable, 0.83 ± 0.06 and 0.68 ± 0.22 , respectively. The large standard deviation of the measurement from the film was due to uncertainty in measuring the reflectance which was complicated by the primary mask pattern and light piping within the film. Similar comparisons of the intrinsic quantum efficiency of powder (QE_p) and thin-film (QE_{tf}) samples have been made for other compositions including $Y_{0.9}Al_{0.05}Eu_{0.05}VO_4$ ($QE_p = 0.78 \pm 0.05$, $QE_{tf} = 0.76 \pm 0.12$) and $Y_{1.95}Eu_{0.05}O_3$ ($QE_p = 0.91 \pm 0.04$, $QE_{tf} = 0.88 \pm 0.09$). In all cases the intrinsic quantum efficiencies of the film and the powder of the same composition were equal to within statistical error. This indicates that the extrinsically measured luminosity and chromaticity obtained in high-throughput screens of thin-film libraries of similar morphology are reliable harbingers of the intrinsic efficiencies of the powder samples.

There exist a vast number of unexplored chemical compositions (and processing pathways) that yield materials with possibly useful properties for a myriad of applications, but little *a priori* theoretical understanding to guide in the selection of specific candidate syntheses. With conventional inorganic methods fewer than 1% of all possible ternary compounds and less than 0.01% of possible quaternary compounds have been synthesized so far¹⁶. Combinatorial chemistry is a means of significantly increasing the rate at which new compositions of matter are synthesized and studied. Furthermore, trends in physical properties and performance over large regions of composition (and processing) space may yield important guidance in developing the fundamental understanding and theories that at present are lacking.

While our paper was under review, Sun *et al.*¹⁷ reported another

combinatorial method for preparing much smaller phosphor libraries, which enabled them to identify some new candidate materials. □

Methods

Powder samples. Europium-doped vanadates of the type $Y_{0.975-n-m}Al_mLa_nEu_{0.025}VO_4$ were prepared by firing the well-mixed stoichiometric oxides for 3 h at 1,000 °C, followed by one regrind and one refiring. The brown powders were then stirred in 2 M NaOH for 20 min at 60 °C (to remove unreacted V_2O_5), followed by vacuum filtration and several water washings.

Library processing. Heating and cooling rates varied from 2 °C min⁻¹ to 10 °C min⁻¹. We used this slow thermal cycle to attempt to reduce the complications of thermal stresses which may develop between the films of varying compositions and thicknesses, and the substrate. In spite of such efforts, the film morphology was found to be variable across a library, which is of concern when characterizing relative extrinsic luminescent properties. The large discovery library (Fig. 1a) was heated initially to 500 °C at a rate of 4 °C min⁻¹. After 2 h at 500 °C, it was heated to 850 °C at the same rate and held for 5 h before cooling at 10 °C min⁻¹ to 100 °C. The triangular vanadate library (Fig. 1c) was heated initially to 650 °C at a rate of 2 °C min⁻¹. After 3 h at 650 °C, it was cooled to 100 °C at 2 °C min⁻¹ and removed for initial screening (not shown). The library was then reheated to 900 °C at a rate of 4 °C min⁻¹ and held for 6 h before cooling at 4 °C min⁻¹. The Eu gradient library, which resulted in the data of Fig. 2, was treated similarly.

Received 16 May; accepted 19 August 1997.

- Bunin, B. A., Plunkett, M. J. & Ellman, J. A. The combinatorial synthesis and chemical and biological evaluation of a 1,4-benzodiazepine library. *Proc. Natl Acad. Sci. USA* **91**, 4708–4712 (1994).
- Xiang, X.-D. *et al.* A combinatorial approach to materials discovery. *Science* **268**, 1738–1740 (1995).
- Briceño, G., Chang, H., Sun, X., Schultz, P. G. & Xiang, X.-D. A class of cobalt oxide magnetoresistive materials discovered with combinatorial synthesis. *Science* **270**, 273–275 (1995).
- DiSalvo, F. J. Solid-State chemistry—A rediscovered chemical frontier. *Science* **247**, 649–655 (1990).
- Maruska, H. P., Parodos, T., Kalkhoran, N. M. & Halverson, W. D. Challenges for flat panel display phosphors. *Mater. Res. Soc. Symp. Proc.* **269**–280 (1994).
- Butler, K. H. *Fluorescent Lamp Phosphors* (Pennsylvania State University Press, University Park, PA, 1980).
- Vecht, A. phosphors I: Powders. *SID Seminar Lecture Notes* **2**, F-2/3 (1996).
- Ropp, R. C. *The Chemistry of Artificial Lighting Devices*, 414–656 (Elsevier, Amsterdam, 1993).
- Blasse, G. & Grabmaier, B. C. *Luminescent Materials* (Springer, Berlin, 1994).
- Ouweltjes, J. L. Luminescence and phosphors. *Mod. Mater.* **5**, 161–257 (1965).
- Ozawa, R. & Toshihinao, A. Fluorescent elements based on aluminum-substituted yttrium vanadate. Japanese Patent No. 660215. *Chem. Abstr.* **72**, 16741 (1969).
- Commission Internationale de L'Éclairage Colorimetry 2nd edn (Publication CIE No. 15.2, Central Bureau of the Commission Internationale de L'Éclairage, Vienna, 1986).
- Yokota, W., Shoji, R. & Kimura, K. Yttrium vanadate phosphor. Japanese Patent No. 45041295 B4 701224. *Chem. Abstr.* **75**, 103573 (1970).
- Thornton, W. A. Quantum efficiency spectra of photoluminescent materials. *J. Electr. Chem. Soc.* **116**, 286–298 (1969).
- Ropp, R. C. *Luminescence and the Solid State: Studies in Inorganic Chemistry* **12** (Elsevier, New York, 1991).
- Rodgers, J. R. & Villars, P. Trends in advanced materials data—Regularities and predictions. *Mater. Res. Soc. Bull.* **18**, 27–29 (1993).
- Sun, X.-D., Gao, C., Wang, J. & Xiang, X.-D. Identification and optimization of advanced phosphors using combinatorial libraries. *Appl. Phys. Lett.* **70**, 3353–3355 (1997).

Acknowledgements. This work would not have been possible without technical assistance from I. Campbell, G. Wallace-Freedman, Y. Wang, P. Wang and J. Wu. We also acknowledge helpful advice from I. Goldwasser, S. Jacobsen and P. G. Schultz.

Correspondence should be addressed to W.H.W. (e-mail: hweinber@symyx.com).

Ordered macroporous materials by emulsion templating

A. Imhof & D. J. Pine

Departments of Chemical Engineering and Materials, University of California, Santa Barbara, California 93106-5080, USA

Ordered macroporous materials with pore diameters comparable to optical wavelengths are predicted to have unique and highly useful optical properties such as photonic bandgaps^{1–3} and optical stop-bands⁴. Tight control over the pore size distribution might also lead to improved macroporous materials (those with pores greater than approximately 50 nm) for application as catalytic surfaces and supports⁵, adsorbents, chromatographic materials,

filters⁶, light-weight structural materials⁷, and thermal, acoustic⁸ and electrical insulators⁹. Although methods exist for producing ordered porous materials with pore diameters less than 10 nm (refs 10, 11), there is no general method for producing such materials with uniform pore sizes at larger length scales. Here we report a new method for producing highly monodisperse macroporous materials with pore sizes ranging from 50 nm to several micrometres. Starting with an emulsion of equally sized droplets (produced through a repeated fractionation procedure¹²), we form macroporous materials of titania, silica and zirconia by using the emulsion droplets as templates around which material is deposited through a sol–gel process¹³. Subsequent drying and heat treatment yields solid materials with spherical pores left behind by the emulsion droplets. These pores are highly ordered, reflecting the self-assembly of the original monodisperse emulsion droplets into a nearly crystalline array¹⁴. We show that the pore size can be accurately controlled, and that the technique should be applicable to a wide variety of metal oxides and even organic polymer gels.

Our basic idea is to use sol–gel processing to deposit an inorganic material (for example, a metal oxide) at the exterior of the droplets in a monodisperse emulsion. The method takes advantage of the fact that the oil droplets are both highly deformable and easily removable. The high deformability allows the inorganic gel to accommodate large shrinkage, which prevents cracking and pulverization during ageing and drying. Furthermore, droplet volume fractions can exceed the close-packing limit of 74%. As emulsions are made of liquids, the droplets are easily removed by evaporation or dissolution after the templating has been accomplished. The versatility of this technique offers advantages over existing routes for the production of macroporous material, which are limited to widely nonuniform⁷ or irregularly shaped pores¹⁵, and therefore do not offer the possibility of ordered porosity. Moreover, other techniques are not readily adapted to produce a wide variety of porous materials.

Sol–gel processes make use of metal alkoxides dissolved in a lower alcohol and hydrolysed by the controlled addition of water. (The alcohol is needed as a solvent because its intermediate polarity makes it a good solvent for both the apolar alkoxide and the polar water.) This produces a sol of nanometre-sized particles of the corresponding metal oxide. Further ageing of the sol at predetermined pH and temperature causes growth and aggregation of the particles resulting in gelation. Drying and heat treatment then produce the metal oxide by removing the solvent and residual organics.

Unfortunately, such a procedure cannot be used to produce most metal oxides in conventional aqueous emulsions. The principal difficulty is that most metal oxides are extremely reactive with water and therefore are incompatible with aqueous emulsions. Another difficulty is that large amounts of alcohol will destroy an emulsion because of its tendency to mix both oil and water. Thus, emulsion templating poses two difficulties: first, to find a stable emulsion in which water is replaced by another polar liquid and second, to find a method to do the sol–gel processing in this polar liquid instead of an alcohol.

In earlier work, we developed several nonaqueous emulsions which are stable against demixing induced by droplet coalescence and against coarsening by Ostwald ripening¹⁶. Of these, we found that oil-in-formamide emulsions are the most suitable candidates. They can be stabilized by symmetric triblock copolymers of the general formula (ethylene oxides)_n–(propylene oxide)_m–(ethylene oxide)_n when $2n/m = 0.3–0.4$ and they can be separated into batches consisting of nearly monodisperse droplets using the fractionation procedure referred to above¹². We also found other nonaqueous emulsions which were not suitable for our purposes because they were destabilized by the sol–gel chemistry.

In our emulsion templating procedure, an oil-in-formamide emulsion is prepared and fractionated to obtain the desired

degree of droplet size uniformity¹². Next, a metal oxide sol is prepared in pure formamide. This sol is made by mixing a chemically modified metal alkoxide with formamide which contains a little water, such that the resulting metal/water ratio is in the range 3–10. This partially hydrolyses the alkoxide so that it becomes soluble in formamide but does not react further. This eliminates the

need for other solvents such as alcohol. At this point, the previously prepared emulsion is dispersed in the sol. The droplet volume fraction is adjusted to the desired porosity by centrifugation. Gelation is induced by adding a small amount of ammonia to increase the pH and takes place over several hours. Before gelation, monodisperse droplets self-assemble into a colloidal crystal when

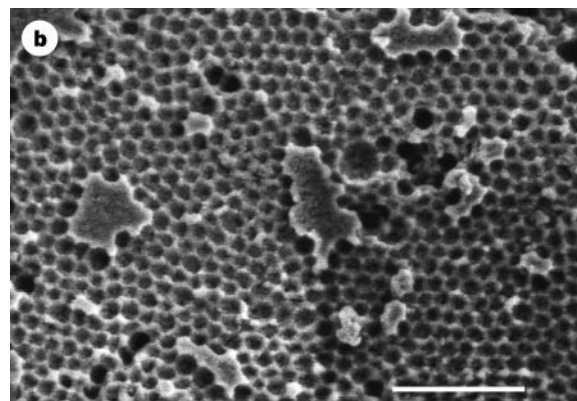
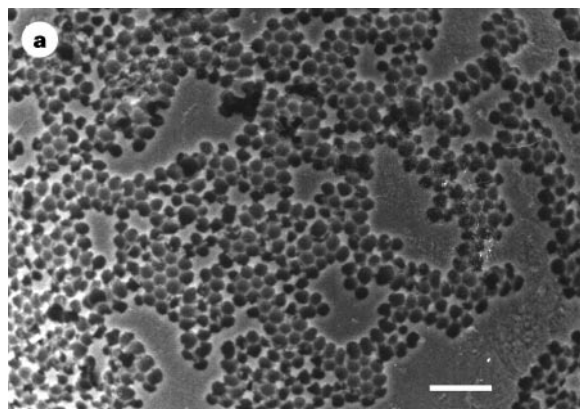


Figure 1 Scanning electron micrographs of porous titania produced in a nearly monodisperse emulsion showing honeycomb structures. The samples had the form of a pellet which was broken to observe the structure inside. **a**, After drying the gel at 60 °C for 2 weeks. The large featureless patches are regions where the pores were not exposed. **b**, The same sample after calcination at 1,000 °C in air for 2 h, showing that the hexagonal pore structure is fully retained. **c**, The structure now consists of small rutile crystallites ~50 nm in size. Scale bars, 1 μm .

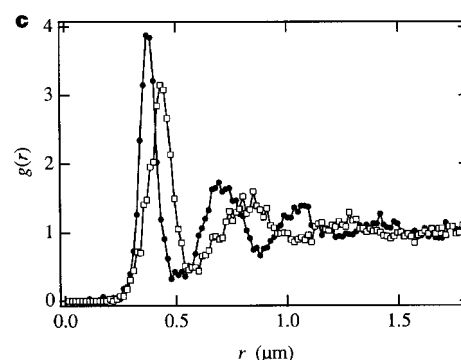
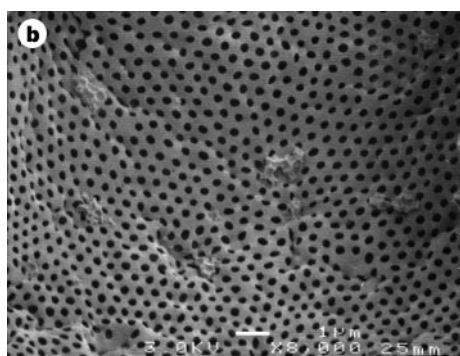
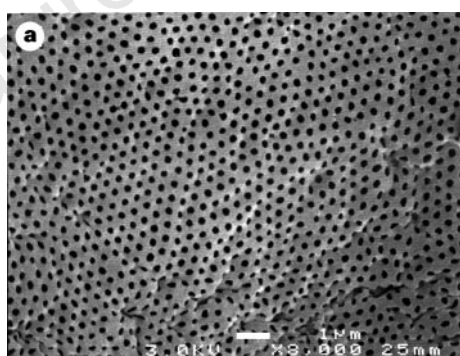
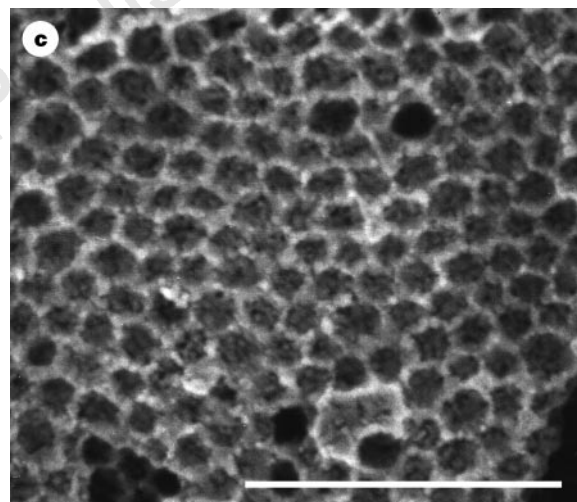


Figure 2 Control over pore size. **a** and **b**, SEM images of porous titania calcined at 1000 °C, prepared from two slightly different emulsion fractions. Although the actual pores are almost touching they do not appear to do so in these images because the surface slices only through the tops of the pores. Scale bars, 1 μm . **c**, Radial distribution functions of the pore centres in the above images, calculated by computer image analysis. They peak at an average inter-pore distance of 0.377 and 0.435 μm , for images **a** (filled circles) and **b** (empty squares) respectively. The half-widths at half-maximum of the first peak are 0.043 and 0.060, respectively. These values are upper limits because the spread in inter-pore distance is increased by the slight buckling of the surfaces seen in the pictures. Note that there are three peaks clearly evident in each of the radial distribution functions (with evidence of a fourth peak) indicating a high degree of local crystalline order.

their volume fraction of droplets exceeds $\sim 50\%$ (ref. 14). The gel is aged to let the reactions go to completion, and then washed in alcohol, dried, and subjected to a heat treatment to remove residual organics.

The pore structure of dried emulsion-templated titania as observed under a scanning electron microscope (SEM) is shown in Fig. 1a. The pores are very uniform with an average diameter of 175 nm and show a high degree of order. The original emulsion had a polydispersity of $\sim 10\%$ and an oil volume fraction of $\sim 55\%$. During the drying stage the linear dimensions of the gel shrank by $\sim 50\%$. Thus, the original droplets had a diameter of ~ 350 nm. After drying, the gel still contains a significant amount of organic material which keeps the metal oxide in an amorphous state. The remaining organics can be removed by heat treatment. SEM images of the same sample after heating at $1,000^\circ\text{C}$ for 2 h are shown in Fig. 1b, c. Remarkably, the pore structure is fully retained, but the pores have shrunk to 145 nm. X-ray diffraction studies of this material showed the spectrum of TiO_2 crystallized in the rutile structure.

We can exercise considerable control over pore size. In Fig. 2a, b we compare SEM images of porous titania made from two slightly different emulsion fractions. The difference in appearance of these pictures from those of Fig. 1 is caused by the fact that the surface 'decapitated' the pores, whereas in Fig. 1 it cut the pores almost exactly in half. By computer image analysis we obtained the radial distribution function $g(r)$ of the pores (Fig. 2c). This function represents the chance of finding the centre of a pore at a distance r from the centre of another pore, relative to that of a completely homogeneous distribution. The high degree of long-range order is indicated by a series of peaks. The average nearest-neighbour distance is given by the location of the first peak, and the width of this peak is a measure for the spread. These results show that we were able to make materials with a well defined pore size difference of $\sim 20\%$.

That the macropore structure is unaffected by the heat treatment is all the more remarkable if one considers the changes that take place on a smaller length scale. Detailed analysis of these changes was made on similarly prepared porous titania starting with a nonuniform emulsion. Samples were heated to different tempera-

tures for 2 h and analysed. X-ray powder diffraction spectra showed that between 400 and 800°C the crystal structure was that of anatase, and that it recrystallized to rutile at higher temperatures. Thermogravimetric analysis showed that 90% of the weight loss takes place below 300°C . Weight loss is complete at 500°C when the gels have lost 30–35% of their initial weight. Nitrogen sorption isotherms were measured to investigate the densification of the titania matrix. From the isotherms, the total internal surface area and the total volume of pores smaller than ~ 50 nm is derived (Fig. 3). The technique is not sensitive to the macropores. The data demonstrate that the titania matrix contained mesopores of 2–10 nm over most of the temperature range (type IV isotherms). Only the sample treated at $1,000^\circ\text{C}$ had a type II isotherm indicative of a nonporous material. Thus, the data indicate strong densification of the titania matrix. Such control over densification is very useful. In applications such as catalysis and sorption, high porosity and high internal surface area improve efficiency. Macropores facilitate material transport to mesoporous internal regions where reactions can take place. By contrast, for optical applications, it is desirable for the matrix to be as dense as possible to achieve maximal contrast in the refractive index between the matrix and the macropores³.

One important advantage of our emulsion templating process is that it can be used to produce porous structures in many different materials. This should allow tailoring of their chemical, electrical, magnetic and optical properties. The choice of materials is illustrated by the SEM images of macroporous zirconia and silica, shown in Fig. 4a and b, respectively. In both cases ordered pores were obtained even after calcination at elevated temperatures.

Templating of nonaqueous emulsions is a versatile and inexpensive way of producing ordered macroporous ceramics of many different materials. By using even more monodisperse emulsions, we should be able to make structures, such as those shown in Figs 1 and 2, with even longer-ranged order. Combination with recently developed techniques for nucleating colloidal crystals on surface microstructures^{17,18} and for preparing binary alloy structures^{19,20}

Figure 3 Effect of calcination temperature on physical properties of porous titania. ▶

Figure shows Brunauer-Emmett-Teller specific surface (circles) and mesoporosity (triangles) of titania produced in a polydisperse emulsion plotted versus calcination temperature (2 h in air, heating rate $15^\circ\text{C min}^{-1}$). The data were obtained from nitrogen sorption isotherms on a Micromeritics ASAP 2000 adsorption instrument using standard procedures. The mesoporosity is the volume fraction of all pores smaller than 50 nm and therefore does not include the macropores. After calcination at 400°C both the high mesoporosity (44%) and the large internal surface ($208\text{ m}^2\text{ g}^{-1}$) indicate that the titania matrix is in itself a very open structure. Calcination at $1,000^\circ\text{C}$ almost completely densifies the matrix.

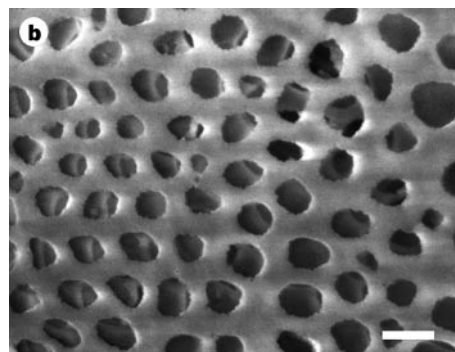
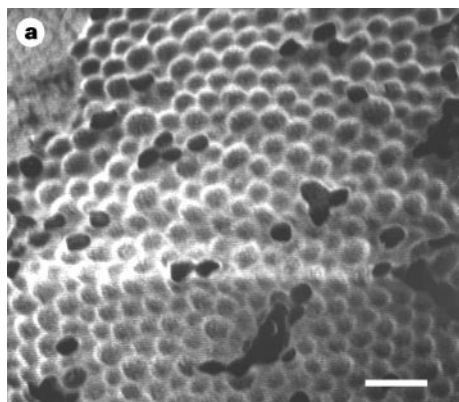
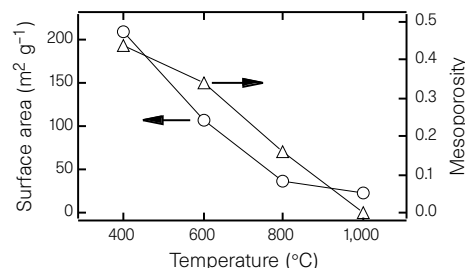


Figure 4 Scanning electron micrographs of emulsion templated materials, showing the applicability of the technique to different materials. **a**, Zirconia with pores of $0.35\text{ }\mu\text{m}$ after heating at $1,000^\circ\text{C}$ in air for 2 h. **b**, Silica with pores of $1.0\text{ }\mu\text{m}$ after heating at 600°C for 2 h. The porosity of this sample is 89%, as calculated

from the density of the material. As this is above the maximum packing fraction of spheres, the droplets have deformed which resulted in a connected set of pores. This allows one to look some way into the porous structure. Scale bars, $1\text{ }\mu\text{m}$.

could bring within reach the production of photonic-bandgap materials^{4–6} at optical wavelengths. □

Methods

Stable nonaqueous emulsions were obtained with formamide (98%) as the polar liquid and isooctane (>99%) as the oil. Before emulsification, the oil was mixed with ~1% of silicone oil to prevent Ostwald ripening¹⁶. The surfactant was a symmetric triblock copolymer, poly(ethylene glycol)–poly(propylene glycol)–poly(ethylene glycol) with a relative molecular mass of 5,800 containing 30% by weight of ethylene glycol monomer (Aldrich). Emulsification was done with a homogenizer. The emulsion was fractionated as described by Bibette¹².

The preparation of the ceramic precursor solution is described with the titania synthesis as an example, because it requires the most reactive alkoxide which is the most difficult to handle. Titanium tetraisopropoxide (97%) was treated with an equimolar amount of the chelating agent 2,4-pentanedione to reduce its reactivity towards water²¹. It was then mixed rapidly by stirring with a mixture of water and formamide, such that the resulting H₂O/Ti molar ratio was 3.5 and the Ti concentration was 2.0 M. The resulting clear yellow solution contains a considerable amount of isopropanol produced by the hydrolysis reaction. Because alcohol will destabilize the emulsion, it was removed by double extraction with a fivefold excess of hexanes. The resulting yellow liquid was often turbid. The sol was heated briefly to ~90 °C producing a clear yellow, slightly viscous sol, which did not form a precipitate for several weeks. Zirconia sols were prepared in an analogous fashion, starting with zirconium tetra-*n*-butoxide (70% in *n*-butanol). Silica sols were prepared by vigorously mixing silicon tetramethoxide with a mixture of water and formamide acidified to pH 2 with HCl.

The surfactant was dissolved to 2 wt% in the sol. For nonuniform pore materials the oil was emulsified directly into it. For uniform pores the separately prepared monodisperse emulsion was centrifuged and the cream dispersed in the sol. The droplet volume fraction was then set by centrifugation and removal of the clear yellow fluid at the bottom. Then 30% ammonia was added so that the molar ratio NH₃/Ti ≈ 1. The increase in pH induces gelation in about 3 h. Gels were aged at 50 °C for 24 h in closed 2-ml polyethylene vials, then allowed to exchange pore fluids with excess ethanol for 24 h. The gels were then dried at room temperature and calcined in air in a furnace.

Received 27 February; accepted 11 August 1997.

1. Yablonovitch, E. Photonic band-gap structures. *J. Opt. Soc. Am. B* **10**, 283–295 (1993).
2. Joannopoulos, J. D., Meade, R. D. & Winn, J. N. *Photonic Crystals: Molding the Flow of Light* (Princeton Univ. Press, 1995).
3. Soukoulis, C. *Photonic Band Gap Materials* (Kluwer, Dordrecht, 1996).
4. Flaugh, P. L., O'Donnell, S. E. & Asher, S. A. Development of a new optical wavelength rejection filter: Demonstration of its utility in Raman spectroscopy. *Appl. Spectrosc.* **38**, 847–850 (1984).
5. Harold, M. P. et al. Catalysis with inorganic membranes. *MRS Bull.* **19**, 34–39 (1994).
6. Bhawe, R. R. *Inorganic membranes synthesis, characteristics, and applications* (Van Nostrand Reinhold, New York, 1991).
7. Wu, M. X., Fujii, T. & Messing, G. L. Synthesis of cellular inorganic materials by foaming sol gels. *J. Non-Cryst. Solids* **121**, 407–412 (1990).
8. Litovsky, E., Shapiro, M. & Shavit, A. Gas pressure and temperature dependences of thermal conductivity of porous ceramic materials. 2. Refractories and ceramics with porosity exceeding 30%. *J. Am. Ceram. Soc.* **79**, 1366–1376 (1996).
9. Singer, P. Low k dielectrics: the search continues. *Semicond. Int.* **19**, 88–96 (1996).
10. Beck, J. S. et al. A new family of mesoporous molecular sieves prepared with liquid crystal templates. *J. Am. Chem. Soc.* **114**, 10834–10843 (1992).
11. Krauss, T., Song, Y. P., Thoms, S., Wilkinson, C. D. W. & DelaRue, R. M. Fabrication of 2-D photonic bandgap structures in GaAs/AlGaAs. *Electron. Lett.* **30**, 1444–1446 (1994).
12. Bibette, J. Depletion interactions and fractionated crystallization for polydisperse emulsion purification. *J. Colloid. Interface Sci.* **147**, 474–478 (1991).
13. Brinker, C. J. & Scherer, G. W. *Sol-Gel Science* (Academic, San Diego, 1990).
14. Pusey, P. N. & van Megen, W. Phase behaviour of concentrated suspensions of hard colloidal spheres. *Nature* **320**, 340–342 (1986).
15. Walsh, D. & Mann, S. Fabrication of hollow porous shells of calcium carbonate from self-organizing media. *Nature* **377**, 320–323 (1995).
16. Imhof, A. & Pine, D. J. Stability of nonaqueous emulsions. *J. Colloid Interface Sci.* (in the press).
17. Dinsmore, A. D., Yodh, A. G. & Pine, D. J. Entropic control of particle motion using passive surface microstructures. *Nature* **383**, 239–242 (1996).
18. van Blaaderen, A., Ruel, R. & Wiltzius, P. Template-directed colloidal crystallization. *Nature* **385**, 321–324 (1997).
19. Hachisu, S. & Yoshimura, S. Optical demonstration of crystalline superstructures in binary mixtures of latex globules. *Nature* **283**, 188–189 (1980).
20. Bartlett, P., Ottewill, R. H. & Pusey, P. N. Superlattice formation in binary mixtures of hard-sphere colloids. *Phys. Rev. Lett.* **68**, 3801–3804 (1992).
21. Sanchez, C., Babonneau, F., Doeuff, S. & Leautaud, A. in *Ultrastructure Processing of Advanced Ceramics* (eds Mackenzie, J. D. & Ulrich, D. R.) 77–87 (Wiley, New York, 1988).

Acknowledgements. We thank F. Lange and G. Stucky for discussions, and Q. Huo for help with the adsorption measurements.

Correspondence should be addressed to D.J.P. (e-mail: pine@engineering.ucsb.edu).

Experimental determination of the organic carbon flux from open-ocean surface waters

S. Emerson*, P. Quay*, D. Karl†, C. Winn†, L. Tupas† & M. Landry†

* School of Oceanography, University of Washington, Seattle, Washington 98195, USA

† Department of Oceanography, University of Hawaii, Honolulu, Hawaii 96822, USA

The flux of biologically produced organic carbon from the euphotic zone of the ocean to the deep waters below—the ‘biological organic carbon pump’—is one of the main controls on the carbon dioxide partial pressure in the atmosphere¹. Accurate determination of this flux is therefore critically important for understanding the global carbon cycle and its response to climate change. Our goal is to assess how accurately the biological organic carbon pump can be determined at a single location and to constrain estimates of the global value. As there are no standards against which such environmental fluxes can be measured, we assess accuracy by comparing results from three independent experimental approaches for measuring the net annual export of organic carbon from the euphotic zone in the subtropical North Pacific Ocean near Hawaii. Mass balances of dissolved oxygen, inorganic carbon and organic carbon yield estimates of the organic carbon export flux of 2.7 ± 1.7 , 1.6 ± 0.9 and $2.0 \pm 0.9 \text{ mol C m}^{-2} \text{ yr}^{-1}$, respectively. These three estimates are not significantly different, and establish the present analytically attainable accuracy at this location to be about ±50%. If $2.0 \text{ mol C m}^{-2} \text{ yr}^{-1}$ is typical of the organic carbon export flux in the subtropical ocean, then this vast region, often considered to be a biological desert, may be responsible for up to half of the global-ocean biological organic carbon pump.

The United States JGOFS (Joint Global Ocean Flux Study) Pacific time-series station² (station ALOHA of the Hawaii Ocean Time series) is located in the southeastern region of the anticyclonic subtropical gyre at 22° 45' N, 150° W. Particulate carbon fluxes have been determined over the duration of the programme using drifting sediment traps³, as have measurements of the dissolved inorganic carbon (DIC) concentration⁴ and its stable-isotope ratio⁵ (¹³C/¹²C). Monthly measurements of the concentrations of the gases O₂, N₂ and Ar were made during the years 1990⁶, 1992 and 1995.

The biological consumption of DIC (J_{DIC} in $\text{mol C m}^{-2} \text{ yr}^{-1}$) in the euphotic zone of the ocean is equal to the net production of particulate inorganic carbon (PIC) as CaCO₃ (J_{PIC}) and organic carbon (J_{OC}) in both dissolved and particulate forms (J_{DOC} and J_{POC} , respectively):

$$J_{\text{DIC}} = J_{\text{PIC}} + J_{\text{OC}} = J_{\text{PIC}} + J_{\text{DOC}} + J_{\text{POC}} \quad (1)$$

Sediment trap collections and alkalinity/DIC ratios in the upper thermocline⁷ at station ALOHA indicate that the flux of PIC out of the upper ocean is only about one-tenth of the total particulate carbon flux; thus, we focus on the net biological production rate of organic carbon, which is stoichiometrically related to the net production of oxygen (J_{O_2}) via the production quotient⁸, β :

$$J_{\text{OC}} = \beta J_{\text{O}_2} \quad (2)$$

Our approach is to evaluate each of the production terms in these equations by determining mass balances of DIC, oxygen and organic carbon.

The mass balance of dissolved species, A, in the upper ocean at this location can be expressed as:

$$h[d(\bar{A}/dt)] = F_{a-w} + F_h - V\Delta\bar{A} \pm J_A \quad (3)$$

The over-bar indicates the mean value over the upper 100 m, which approximates the depth of the winter mixed layer, h , and the transition from net community O_2 production to net community respiration⁶; F_{a-w} is the gas exchange flux, and F_h is the flux between the mixed layer and the top of the thermocline. The latter value includes the Ekman downwelling and all other 'vertical' transport processes. Surface horizontal advection is modelled as a flow rate, V (in cm s^{-1}), multiplied by the difference in dissolved concentration between the upstream and local concentrations, $\Delta\bar{A}$. The net biological production or consumption term is indicated by J_A .

Net biological oxygen production is determined from dissolved O_2 , N_2 and Ar concentrations measured at roughly monthly intervals for the years 1990, 1992 and 1995 (Fig. 1). The positive difference between the degree of oxygen and argon supersaturation (Fig. 1a) is a qualitative measure of the biological oxygen production⁶ in the euphotic zone. The gases are supersaturated in surface waters during the entire year (Fig. 1b), indicating a net diffusive transport to the atmosphere. Air–water exchange rates, F_{a-w} , are modelled as the sum of a bubble flux term, B , and a diffusion term^{9–11} which is the product of a mass transfer coefficient, G (in m d^{-1}), and the degree of gas saturation:

$$F_{a-w} = B + G[S(p_C) - [C]] \quad (4)$$

where p_C is the partial pressure of gas C (in atm) in the atmosphere, S is the gas solubility (in $\text{mol m}^{-3} \text{atm}^{-1}$), and $[C]$ is the gas concentration (in mol m^{-3}) in the mixed layer. For relatively insoluble gases, the bubble term is influential in causing gas supersaturation (see Fig. 1 legend). The mass transfer coefficient is evaluated from daily averaged wind speed measurements on a nearby buoy⁶ and correlations to wind speed^{12,13} (mean annual value, 2.7 m d^{-1}). Horizontal velocities, V , are determined to be $1\text{--}2 \text{ cm s}^{-1}$ from the northward component of climatological ship drift speeds and

geostrophic calculations¹⁴. With these independent transport estimates, three simultaneous equations for N_2 , Ar and O_2 are used to calculate O_2 fluxes via bubble transport, B , vertical exchange between the surface ocean and thermocline, F_h , and the net biological oxygen production rate, J_{O_2} . The mass balance described in equation (3) is solved using a three-layer, two-dimensional model that separates the mixed layer, the mixed layer to 100 m depth, and the upper thermocline (see ref. 6 for details of the model.)

The oxygen mass balance indicates that net biological O_2 production in the euphotic zone is roughly balanced by losses to the atmosphere and the thermocline. The annual rate of change of O_2 concentration and the horizontal flow of O_2 are less than one-tenth the value of the other terms. Oxygen loss to the atmosphere averages $-1.6 \text{ mol O}_2 \text{ m}^{-2} \text{ yr}^{-1}$ for the 3 years. This is the net gas exchange value which consists of the difference between outgassing driven by the surface ocean supersaturation (Fig. 1b) at a rate of $-2.6 \text{ mol m}^{-2} \text{ yr}^{-1}$ and the bubble-induced invasion rate of $1.0 \text{ mol m}^{-2} \text{ yr}^{-1}$. Downward transport to the top of the thermocline is $-1.5 \text{ mol m}^{-2} \text{ yr}^{-1}$ which probably results from turbulent mixing or eddy-induced dipycnal transport¹⁵, as Ekman velocities¹⁴ can account for only a small fraction of this flux. Oxygen losses from the upper ocean are balanced by a net biological oxygen production of $3.6 \pm 2.2 \text{ mol m}^{-2} \text{ yr}^{-1}$, which is equivalent to an organic carbon production rate of $2.7 \pm 1.7 \text{ mol C m}^{-2} \text{ yr}^{-1}$. The conversion from oxygen fluxes to carbon fluxes and error estimates are described in Table 1.

Dissolved inorganic carbon and carbon isotope (DI^{13}C) measurements at station ALOHA are used to determine the net biological carbon export^{5,6}. We derive independent information from these two budgets because DI^{13}C has an equilibration time with the atmosphere via gas exchange that is approximately ten times that of $\text{DIC}^{16,17}$. DIC budgets are constructed for February–November 1990, and February 1994 to February 1995 using the data in Fig. 2 and equations (3) and (4). Because CO_2 is a relatively soluble gas, the air–water exchange process by bubbles is relatively small and has been neglected¹⁰. The gas exchange mass transfer coefficient, G_{CO_2} ,

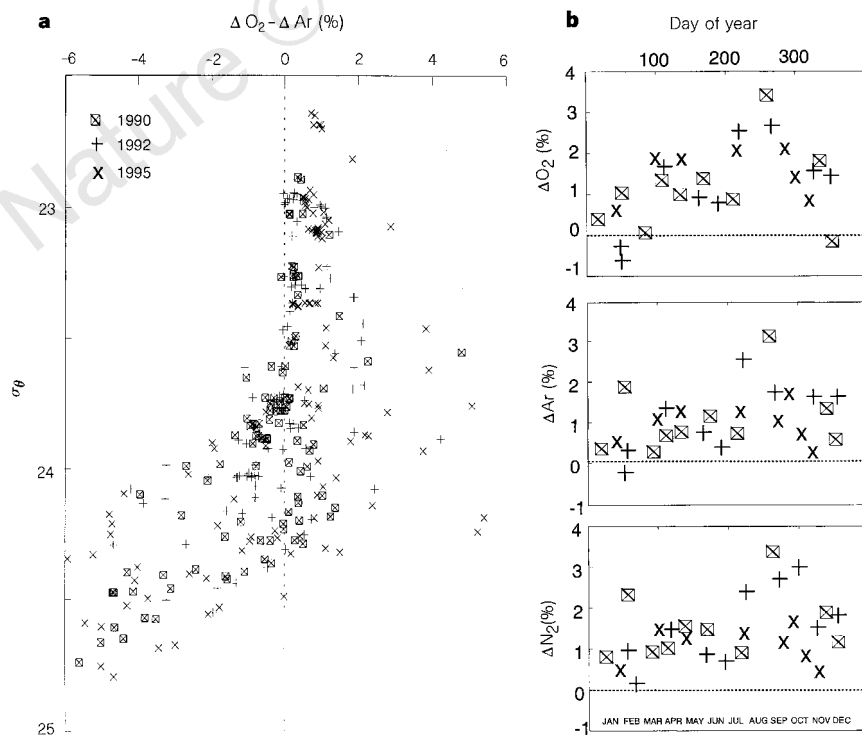


Figure 1 Oxygen, argon and nitrogen data in the upper water column at station ALOHA. Symbols represent different years: 1990 ($\square$), 1992 (+), 1995 ($\times$). **a**, The difference in the degree of saturation, Δ (in %), between oxygen and argon as a function of potential density, σ_θ , indicating a maximum in net biological oxygen production beneath the mixed layer (100 m is about $\sigma_\theta = 24$). **b**, The degree of saturation in the mixed layer as a function of time during the three sampling periods. Each point represents the mean of 3–10 measurements. We note that the waters are supersaturated during the entire year in spite of the diurnal temperature change of 4°C at this latitude. Oxygen was determined by Winkler titration and other gases by ratios to oxygen measured by mass spectrometry (see ref. 28). For the relatively insoluble gases O_2 , N_2 and Ar, the bubble term (B in equation (4)) is important in determining the gas supersaturation. As the mechanism of bubble process cannot be evaluated with only two inert-gas tracers, and there is evidence that in the subtropics simple air injection (where B is proportional to the partial pressure in the atmosphere, p_C) is not adequate^{11,15}, we assume that the mean value for the bubble term is determined by an equal combination of air injection and exchange across the surface of larger bubbles^{9,11} (where B is proportional to the molecular diffusion coefficient (D), solubility (S) and mole fraction p_C , $D^{0.7}S^{0.35}p_C$). The error estimate of $\pm 25\%$ given in Table 1 spans the range between these two extremes.

and the horizontal velocity, V , were estimated as for O_2 , Ar and N_2 . With these values determined independently, we solved two equations for DIC and $DI^{13}C$ to determine the flux between the mixed layer and thermocline, F_h , and the net biological carbon production, J_{OC} . Details of this approach are described in ref. 16.

Measured annual rates of change of DIC and $DI^{13}C$ concentrations were of the same magnitude as their measurement errors of $\pm 2 \mu\text{mol kg}^{-1}$ and $\pm 0.2\%$, respectively. Combination of the mass transfer coefficient and the observed degree of CO_2 gas undersaturation (Fig. 2b) resulted in a net CO_2 invasion rate of $0.6 \text{ mol C m}^{-2} \text{ yr}^{-1}$. A meridional surface DIC flux of $-0.3 \text{ mol C m}^{-2} \text{ yr}^{-1}$ was calculated from the horizontal velocity, V , and measured surface DIC gradients that increase to the north at a rate of $1 \mu\text{mol kg}^{-1}$ per degree of latitude (P.Q., unpublished data). Rates of organic carbon production and flux from the thermocline necessary to balance the DIC fluxes and accommodate the stable-isotope data (Fig. 2) were 1.8 ± 0.9 and $1.5 \pm 1.0 \text{ mol C m}^{-2} \text{ yr}^{-1}$, respectively. The origin of the error estimates is described in Table 1. Both the DIC and oxygen mass-balances indicate that net biological carbon and oxygen production at this location of the ocean are balanced by roughly equal fluxes between the atmosphere–water and mixed layer–thermocline interfaces.

The organic carbon mass balance consists of both particulate (POC) and dissolved (DOC) components. The rate of change of POC concentration is equal to the net formation rate of POC, J_{POC} , minus the fluxes out of the euphotic zone via gravitational settling, F_{POC} , and the net removal of POC by daily zooplankton migrations, F_{ZOO} (in $\text{mol m}^{-2} \text{ yr}^{-1}$):

$$h(d\text{POC}/dt) = J_{POC} - F_{POC} - F_{ZOO} \quad (5)$$

Particulate carbon fluxes measured by floating sediment traps³ at 150 m (Fig. 3) vary between 0.4 and $0.9 \text{ mol C m}^{-2} \text{ yr}^{-1}$ for the years in which other measurements were made. To compare these values

Table 1 Net annual production of organic matter in the upper 100 m of the ocean at station ALOHA

Year	O_2, N_2, Ar	DIC/ $DI^{13}C$	Organic carbon			ΣOC
			POC	DOC	ZOO	
1990	1.7 ± 1.4	1.5 ± 0.8	1.3 ± 0.4	0.8 ± 0.9		2.3 ± 1.0
1992	3.1 ± 2.2		0.8 ± 0.2	1.3 ± 1.2		2.3 ± 1.2
1994		1.8 ± 1.1			0.2 ± 0.1	
1995	3.1 ± 1.2		0.8 ± 0.1	0.5 ± 0.4	0.2 ± 0.1	1.5 ± 0.4
Mean	2.7 ± 1.7	1.6 ± 0.9				2.0 ± 0.9

Values (in $\text{mol C m}^{-2} \text{ yr}^{-1}$) were determined by: mass balance of O_2 , Ar and N_2 (column 2); mass balance of DIC and $DI^{13}C$ (column 3); and the sum (column 7) of fluxes of particulate organic carbon (POC; column 4), dissolved organic carbon (DOC; column 5), and zooplankton (ZOO; column 6) transport. Net oxygen production calculated from the model was transformed to carbon using a production quotient (PQ) of 1.25 which assumes that half of the net nitrogen uptake is as ammonia (PQ = 1.1) and half as nitrate (PQ = 1.4)²⁵. For comparison with other fluxes, the POC flux at 150 m was multiplied by 0.9 to convert it to organic carbon flux⁷ and extrapolated to 100 m using trap flux versus depth regressions established from similar sediment-trap studies in the subtropical Pacific Ocean that included the 100-m depth horizon²³. The extrapolation increases the POC flux by 35%. Uncertainty estimates presented as ($\pm$) were determined in the oxygen and DIC mass balances using a Monte Carlo method in which deviations of the most important input values into equation (3) are assumed to be distributed in a gaussian way about the mean, with standard deviations equal to the uncertainty estimates. Values are chosen randomly and independently from these distributions and the equation solved repeatedly to determine the error or the dependent variables^{6,26}. For the O_2 , Ar and N_2 mass balance, input uncertainties (standard deviations) were: $\pm 30\%$ for the gas exchange parameter, G (this represents the difference between values generated using different correlations with wind speed^{22,23} at 7 m s^{-1} , the mean value at station ALOHA); $\pm 0.2\%$ for the concentration of each of the gases and its degree of saturation⁶; and $\pm 25\%$ in the value of the bubble flux (this spans the range of values estimated by the two different mechanisms explained in Fig. 1 legend). For the DIC and $DI^{13}C$ mass balance, errors in the input values are the same for the gas exchange mass transfer coefficient, G , ($\pm 30\%$); and $\pm 25\%$ for the p_{CO_2} difference between the air and water. DIC and the $\delta^{13}C$ of DIC measurements have errors of $\pm 2 \mu\text{mol kg}^{-1}$ and $\pm 0.02\%$, respectively. Organic carbon is assumed to have a $\delta^{13}C$ mean value and standard deviation of $-21 \pm 2\%$ (ref. 27). The error estimates for organic carbon fluxes represent: (1) for POC, the standard deviation of the monthly mean flux measurements ($n = 9-11$ for each year); (2) for DOC, combination of the error estimates of the oxygen flux, F_h , determined from the O_2 , N_2 and Ar mass balances, and the standard deviation of the linear relationship between the DOC and O_2 in the 75–150 m interval ($\Delta\text{DOC}/\Delta O_2 = 0.58 \pm 0.37$); and (3) for ZOO, the error estimate of $\pm 50\%$ is from the standard deviation of the migratory of zooplankton biomass during monthly cruises in a given year. Mean values are unweighted averages of both the annual fluxes and their standard deviations. A random gaussian-weighted selection of values determined from the oxygen (2.7 ± 1.7) and DIC (1.6 ± 0.9) fluxes indicates that these distributions are indistinguishable $\sim 50\%$ of the time.

with export fluxes determined by O_2 and DIC mass balances, we converted particulate carbon fluxes to organic carbon fluxes and extrapolated to 100 m, as described in Table 1.

Zooplankton respiration at depth represents an organic-carbon transfer out of the upper ocean for species that feed in the euphotic

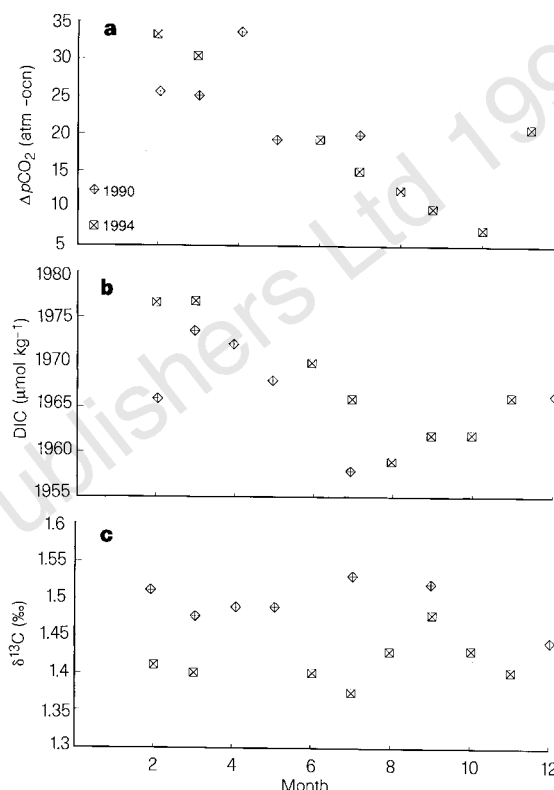


Figure 2 Measurements of dissolved inorganic carbon (DIC) species in the mixed layer at station ALOHA during 1990 (◇) and 1994 (□). **a**, Δp_{CO_2} (p.p.m.) (defined as $p_{CO_2}(\text{atm}) - p_{CO_2}(\text{surface ocean})$; here surface ocean p_{CO_2} values are calculated in 1990³ and measured in 1994 (T. Takahashi, personal communication). **b**, DIC normalized to 35‰ salinity. **c**, $\delta^{13}C$ of DIC (‰ versus PDB). Details of the analytical methods used are given in refs 4 and 16. The procedure for writing the mass balances, including all important isotope ratios and pertinent isotope fractionation factors, are given in ref. 16.

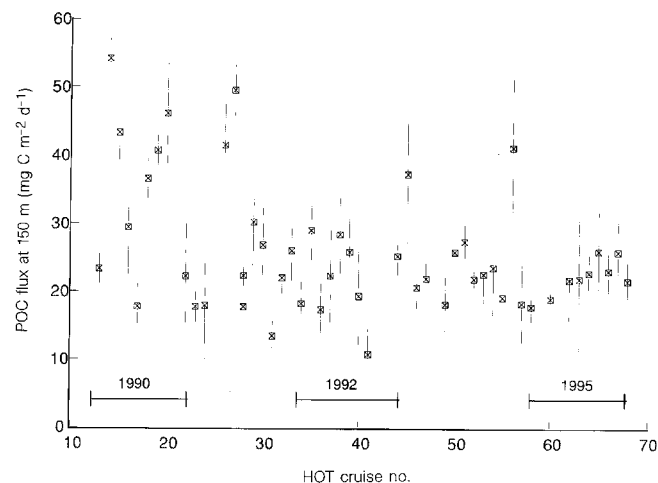


Figure 3 The particulate carbon flux at the depth of 150 m as a function of the Hawaii Ocean Time series (HOT) cruise number, with the 1990, 1992 and 1995 intervals indicated. Vertical error bars indicate the standard deviation of fluxes determined in 2–4 separate sediment-trap collector tubes at the 150-m reference depth deployed for 2–3 days on a free-drifting trap array. Methods are described in ref. 3.

zone at night and migrate to deeper waters during the day¹⁸. Based on measured day–night differences of zooplankton biomass, the size structure of the migrating zooplankton community (M.L., unpublished results), and empirical relationships for zooplankton metabolic rates as a function of animal size and environment temperature¹⁹, we estimate the migrating zooplankton flux to be $0.2 \pm 0.1 \text{ mol C m}^{-2} \text{ yr}^{-1}$.

The strong vertical gradient in DOC in the upper 200 m of the subtropical ocean^{20,21} suggests net production in the euphotic zone and net consumption in the upper thermocline. We estimate the importance of the flux of DOC to the organic carbon balance by comparing the vertical gradients of DOC and oxygen and multiplying the $\Delta\text{DOC}/\Delta\text{O}_2$ ratio by the flux of oxygen across the 100-m depth horizon, F_{100} . This portion of the organic-matter flux is, thus, dependent on results of the oxygen mass balance. The $\Delta\text{DOC}/\Delta\text{O}_2$ ratio across the 100-m horizon (~ 75 – 125 m) for nine different profiles in the years 1993–95 averaged 0.58 ± 0.37 (L.T., unpublished results) providing the driving force for $\sim 40\%$ of the total carbon flux out of the euphotic zone (Table 1).

Three organic-matter fluxes for the years 1990, 1992 and 1995 result in a mean of $2.0 \pm 0.9 \text{ mol C m}^{-2} \text{ yr}^{-1}$. This value is similar to that determined using the DIC and DI^{13}C mass balance (1.6 ± 0.9) and somewhat smaller than the estimate from oxygen mass balance (2.7 ± 1.7). The two most disparate of these three fluxes, however, are not significantly different (Table 1). We suggest that the present attainable accuracy of the biological pump at this location is roughly $\pm 50\%$, and is limited primarily by uncertainty estimates of the individual fluxes. This may not be true for all ocean areas, however, as a recent comparison between organic carbon export determined by oxygen mass balance and POC + DOC fluxes at the Bermuda time-series station yielded a factor of two difference, with the latter being smaller²⁰.

We assume that our estimate of the organic-carbon export at station ALOHA ($\sim 2 \text{ mol C m}^{-2} \text{ yr}^{-1}$) is typical of the subtropical oceans, as it is probably more representative than the station at Bermuda with its very deep winter mixed layers. As the area of the subtropical oceans is $\sim 60\%$ of the total ocean area²², the annual biological pump from this ocean province is 5 – 6 Gt C yr^{-1} . Current estimates of the global ocean organic carbon pump based on sediment trap fluxes are $\sim 5 \text{ Gt C yr}^{-1}$ ($1.2 \text{ mol C m}^{-2} \text{ yr}^{-1}$)^{3,22,23}. Our value of 5 – 6 Gt C yr^{-1} for the subtropics alone indicates that these global estimates are almost certainly too low, as upwelling and coastal regions are believed to account for a significant organic-carbon export²². If we assume that recent model-derived estimates of the global ocean organic carbon pump (10 – 11 Gt yr^{-1})^{1,24} are correct, then the subtropical oceans may account for up to half of the global value. An urgent task for marine scientists is to understand the mechanisms that limit this organic-carbon export so that predictions of the response to changes in forcing can be made. □

Received 1 July; accepted 9 September 1997.

- Sarmiento, J. L. & Siegenthaler, U. in *Primary Productivity and Biogeochemical Cycles in the Sea* (eds Falkowski, P. G. & Woodhead, A. D.) 317–332 (Plenum, New York, 1992).
- Karl, D. M. & Lukas, R. The Hawaii Ocean Time-series (HOT) program: Background, rationale and field implementation. *Deep-Sea Res.* **43**, 129–147 (1996).
- Karl, D. M. *et al.* Seasonal and interannual variability in primary production and particle flux at station ALOHA. *Deep-Sea Res.* **43**, 539–569 (1996).
- Winn, C. D., Mackenzie, F. T., Carrillo, C. J., Sabine, C. L. & Karl, D. M. Air–sea carbon dioxide exchange in the North Pacific subtropical gyre: Implications for the global carbon budget. *Glob. Biogeochem. Cycles* **8**, 157–163 (1994).
- Quay, P. D. & Andersen, H. Organic carbon export rates in the Subtropical N. Pacific. *Eos* **76**, OS85 (1996).
- Emerson, S., Quay, P. D., Stump, C., Wilber, D. & Schudlich, R. Chemical tracers of productivity and respiration in the subtropical Pacific Ocean. *J. Geophys. Res.* **100**, 15873–15887 (1995).
- Sabine, C. L., Mackenzie, F. T., Winn, C. & Karl, D. M. The influence of air–sea exchange on the isotopic composition of oceanic carbon: Observations and modelling. *Glob. Biogeochem. Cycles* **9**, 637–652 (1995).
- Laws, E. A. Photosynthetic quotients, new production and net community production in the open ocean. *Deep-Sea Res.* **38**, 143–167 (1991).
- Keeling, R. On the role of large bubbles in air–sea exchange and supersaturation in the ocean. *J. Mar. Res.* **51**, 237–271 (1993).
- Woolf, D. Bubbles and air–sea transfer velocity of gases. *Atmos. Ocean* **31**, 517–540 (1993).
- Schudlich, R. & Emerson, S. Gas saturation in the surface ocean: The roles of heat flux, gas exchange and bubbles. *Deep-Sea Res.* **43**, 569–591 (1996).

- Liss, P. S. & Merlivat, L. in *The Role of Air–Sea Exchange in Geochemical Cycling* (ed. Buat-Ménard, P.) 112–127 (Reidel, Norwell, MA, 1986).
- Wanninkhof, R. Relationship between wind speed and gas exchange over the ocean. *J. Geophys. Res.* **97**, 7373–7382 (1992).
- Roden, G. I. On the subtropical frontal zone north of Hawaii in winter. *J. Phys. Oceanogr.* **10**, 342–362 (1980).
- Jenkins, W. Using anthropogenic tritium and ^3He to study subtropical gyre ventilation and circulation. *Phil. Trans. R. Soc. Lond. A* **325**, 43–61 (1988).
- Zhang, J. & Quay, P. D. The total organic carbon export rate based on ^{13}C and ^{12}C of DIC budgets in the Equatorial Pacific region. *Deep-Sea Res.* **II** (in the press).
- Broecker, W. S. & Peng, T. H. Gas exchange rates between air and sea. *Tellus* **XXVI**, 21–35 (1974).
- Longhurst, A. R. & Harrison, W. G. The biological pump: profiles of plankton and consumption in the upper ocean. *Prog. Oceanogr.* **22**, 46–123 (1989).
- Iked, T. Metabolic rates of epipelagic marine zooplankton as a function of body mass and temperature. *Mar. Biol.* **85**, 1–11 (1985).
- Carlson, C. A., Ducklow, H. W. & Michaels, A. F. Annual flux of dissolved organic carbon from the euphotic zone of the northwest Sargasso Sea. *Nature* **371**, 405–408 (1994).
- Tupas, L. M., Popp, B. N. & Karl, D. Dissolved organic carbon in oligotrophic waters: Experiments on sample preservation, storage and analysis. *Mar. Chem.* **41**, 207–216 (1993).
- Eppley, R. W. & Peterson, B. J. Particulate organic flux and planktonic new production in the deep ocean. *Nature* **282**, 677–680 (1979).
- Martin, J. H., Knauer, G. A., Karl, D. M. & Broenkow, B. B. VERTEX: Carbon cycling in the northeast Pacific. *Deep-Sea Res.* **34**, 267–285 (1987).
- Six, K. D. & Maier-Reimer, E. Effect of plankton dynamics on the seasonal carbon fluxes in a ocean general circulation model. *Glob. Biogeochem. Cycles* **10**, 559–583 (1996).
- Karl, D. M. *et al.* The role of dinitrogen fixation in the nitrogen budget of the subtropical North Pacific Ocean. *Nature* **388**, 533–538 (1997).
- Quay, P. D., Emerson, S., Wilber, D. O. & Stump, C. The $\delta^{18}\text{O}$ of dissolved O_2 in the surface waters of the subarctic Pacific: a tracer of biological productivity. *J. Geophys. Res.* **98**, 8447–8458 (1993).
- Goericke, R. & Fry, B. Variations of marine plankton $\delta^{13}\text{C}$ with latitude, temperature and dissolved CO_2 in the world ocean. *Glob. Biogeochem. Cycles* **8**, 85–90 (1994).
- Emerson, S., Quay, P. D., Stump, C., Wilber, D. & Knox, M. O_2 , Ar , N_2 and ^{222}Rn in surface waters of the subarctic Pacific Ocean. *Glob. Biogeochem. Cycles* **5**, 49–69 (1991).

Acknowledgements. We thank C. Stump, D. Wilber and D. Hebel for technical work which helped make this study possible. This work was supported by the NSF.

Correspondence and requests for materials should be addressed to S.E. (e-mail: emerson@u.washington.edu).

Ice–volcano interaction of the 1996 Gjalp subglacial eruption, Vatnajökull, Iceland

Magnús T. Gudmundsson*, Freysteinn Sigmundsson† & Helgi Björnsson*

* Science Institute, University of Iceland, Dunhaga 3, 107 Reykjavik, Iceland

† Nordic Volcanological Institute, University of Iceland, Grensasvegur 50, 108 Reykjavik, Iceland

Volcanic eruptions under glaciers can cause dangerous floods and lahars^{1–3} and create hyaloclastite (fragmented glassy rock) mountains^{4–8}. But processes such as the rate of heat transfer between ice and magma, edifice formation, and the response of the surrounding glacier are poorly understood, because of the lack of data. Here we present observations from the fissure eruption at Vatnajökull ice cap, Iceland, in October 1996. In the 13 days of the eruption 3 km^3 of ice were melted and the erupted magma fragmented into glass forming a hyaloclastite ridge 6 – 7 km long and 200 – 300 m high under 500 – 750 m of ice. Meltwater of temperatures of 15 – 20°C flowed along a narrow channel at the glacier bed into the Grímsvötn subglacial lake for five weeks, before draining in a sudden flood, or jökulhlaup. Subsidence and crevassing of the ice cap occurred over the eruptive fissure and the meltwater path, whereas elsewhere the glacier surface remained intact, suggesting that subglacial eruptions do not trigger widespread basal sliding in warm-based glaciers.

Despite the lack of direct observations of subglacial eruptions, qualitative models of the formation of hyaloclastite mountains within glaciers were presented in the 1940s and 1950s^{5–7} and have been advanced by more recent work^{8–10}. These models are based on geological mapping and involve melting of the overlying ice and the emergence of a mountain within a water-filled depression confined by the surrounding glacier. The lowermost parts of the mountains are often made of pillow lavas, whereas the upper parts are

hyaloclastites. Subaerially erupted lavas may cap the mountain, forming a table mountain or tuya. Mountains considered to have been formed in this way are found in Antarctica⁴, British Columbia⁵ and Iceland^{6–8}. Magma fragmentation into pyroclastic glass and its subsequent alteration into consolidated hyaloclastite was observed during the formation of a marine table mountain in the Surtsey eruption off the south coast of Iceland in 1963–67 (refs 11–13).

The eruption in 1996 took place within Vatnajökull, which is temperate (warm based) and the largest ice cap in Europe. On 30 September, at about 22:00 GMT, the beginning of the eruption was marked by the onset of continuous seismic tremor¹⁴. The subglacial eruptive site was located by aeroplanes the following day, midway between the Bárðarbunga and Grímsvötn volcanoes

(Fig. 1). Heavily crevassed ice cauldrons had formed overnight; no signs of unusual melting were observed before the eruption. A shallow linear depression had developed on the ice surface from the eruptive site towards the Grímsvötn lake, and new fractures on the ice shelf covering the lake indicated increased water level. From the first day of the eruption we monitored the eruptive activity and melting of ice, by mapping changes on the ice cap surface, and by monitoring the accumulation of melt water in the Grímsvötn lake (Figs 2 and 3).

On 13 October when the eruption ended, 3 km³ of ice had melted, with a further 1.2 km³ being melted in the following three months (Fig. 3). Ice cauldron formation was fastest during the first four days of the eruption when melting rate averaged 0.4–0.6 km³ per day. At

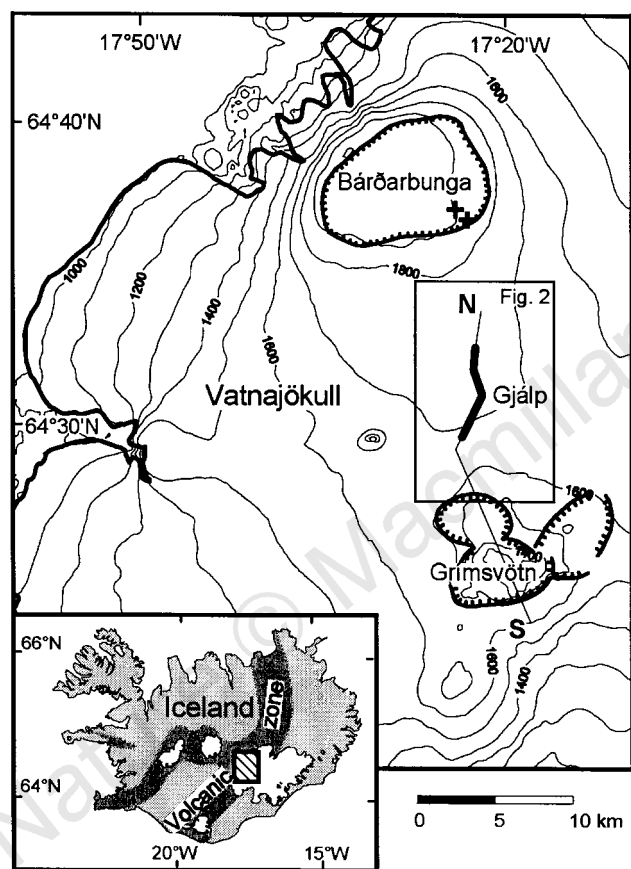


Figure 1 Map of the northwest part of Vatnajökull showing the location of the 1996 eruptive fissure (solid line, in box), midway between the subglacial calderas of the Bárðarbunga and Grímsvötn volcanoes^{22,26}. The contours show ice surface topography (in metres). The eruptive site is well known from accurate mapping and radio-echo soundings that have revealed the bedrock topography²². The Bárðarbunga and Grímsvötn volcanoes are among the most active in Iceland with frequent eruptions and associated jökulhlaups documented in the historical record. A subglacial lake is sustained by geothermal activity within the Grímsvötn caldera^{22,27}. Jökulhlaups originate from the lake, usually at intervals of a few years, but eruptions may disrupt the periodicity of the jökulhlaups^{22,23,28}. The most recent eruptions in the Grímsvötn area occurred in 1934, 1938, 1983 and 1984 (ref. 23). The crosses in the southeast corner of Bárðarbunga are small ice cauldrons considered formed by a minor subglacial eruption at the same time as the main eruption. The locations of the maps and sections in Fig. 2 are shown. Inset shows ice caps and the volcanic zones of Iceland.

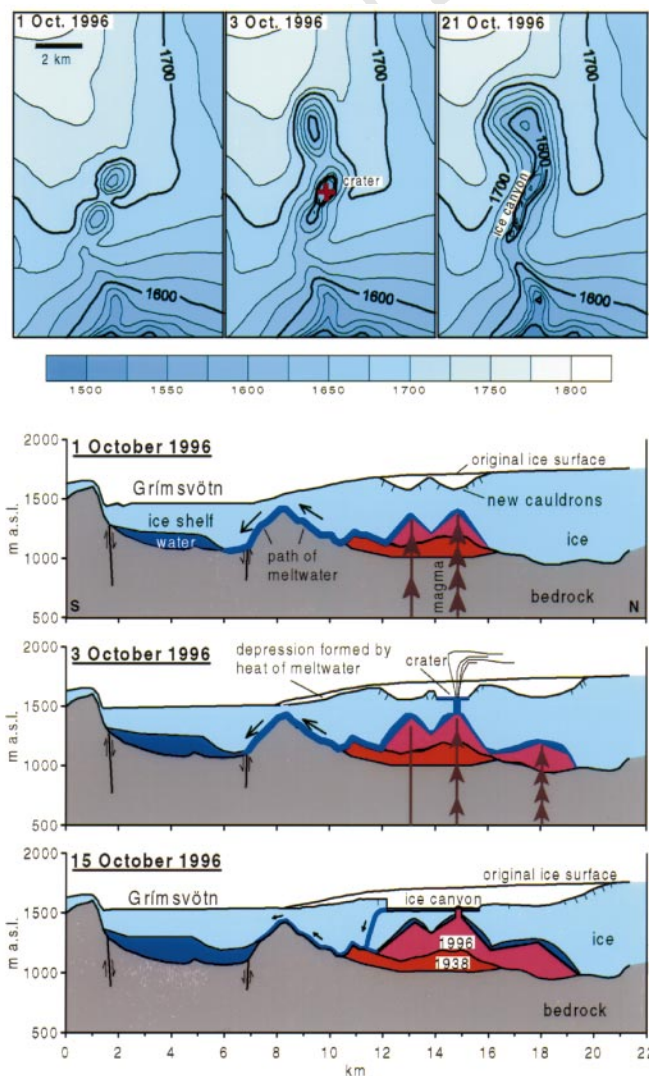


Figure 2 Top, development of ice cauldrons at the eruption site. The cauldrons over the initially 4-km-long eruptive fissure had reached a depth of about 100 m in the first 15 hours. On 2 October the fissure reached a length of 6–7 km as it grew towards the north. Cauldron development was most rapid during the eruption (30 September to 13 October) but continued after the eruption ended. Bottom, sections showing the ice surface and bedrock topography, and, schematically, the formation of a subglacial hyaloclastite ridge at the eruption site, the subglacial meltwater flow path that is mostly determined by ice surface slope, and water accumulation in the Grímsvötn lake (m.a.s.l., metres above sea level). During the latter part of the eruption, a 3.5-km-long canyon developed in the ice surface above the southern part of the volcanic fissure. Water flowed along the canyon, disappearing down into the glacier at its southern end. The material erupted in 1996 lies partly on top of material from an eruption in 1938.

one location the eruption melted its way through 500-m-thick ice in 30 hours. Other parts of the fissure, where initial ice thickness was 550–750 m, remained subglacial. The maximum elevation of the eruption column was 9 km on 3 October, but typically its elevation was 4–5 km and aerially dispersed tephra amounts to no more than 1–2% of the total volume erupted. The visible subaerial eruption was thus always a minor part of the activity. The volcanic fissure has been given the name Gjálþ (after a giantess in Nordic mythology) at the suggestion of G. Larsen.

Formation of ice cauldrons on the surface of ice caps during subglacial eruptions is the response to melting at their base¹⁵ as eruptive products give off thermal and kinetic energy to the environment. If the eruption breaks the ice cover, some of the energy creates an eruption plume. The kinetic energy of eruptive products from typical subaerial fissure eruptions is minuscule when compared with the thermal energy. The total energy of such an eruption is therefore about equal to its thermal energy^{16,17}. Heat released to the environment from eruptive products, Q_e , is

$$Q_e = M_e(C_e\Delta T_e + L_e) \quad (1)$$

where M_e is mass of erupted materials, C_e is their heat capacity, ΔT_e is their drop in temperature and L_e is their latent heat of fusion. Magma erupted subglacially is cooled rapidly and may solidify in a mostly crystalline state as pillow lavas, or, if subjected to lower confining pressure, may fragment and solidify as pyroclastic glass, releasing no latent heat¹⁸, $L_e = 0$. The very fast rate of heat transfer (Fig. 3b) in the Vatnajökull eruption cannot be explained by cooling of pillow lavas, as the solidification time of pillows¹⁹ is at least an order of magnitude too long to account for the observed energy flux per unit area. Thus, quenching and fragmentation of the magma must have been the dominant form of activity. The magma erupted is basaltic andesite with an eruption temperature of $1,090 \pm 50^\circ\text{C}$ and density of $2,700\text{ kg m}^{-3}$ (N. Óskarsson and K. Grönvold, personal communication), and the mean heat capacity of volcanic glasses when quenched from melting temperature to 0°C is about $1.10 \pm 0.05\text{ kJ kg}^{-1}\text{ K}^{-1}$ (ref. 20). If all heat from eruptive products melts ice that is at the melting point, then

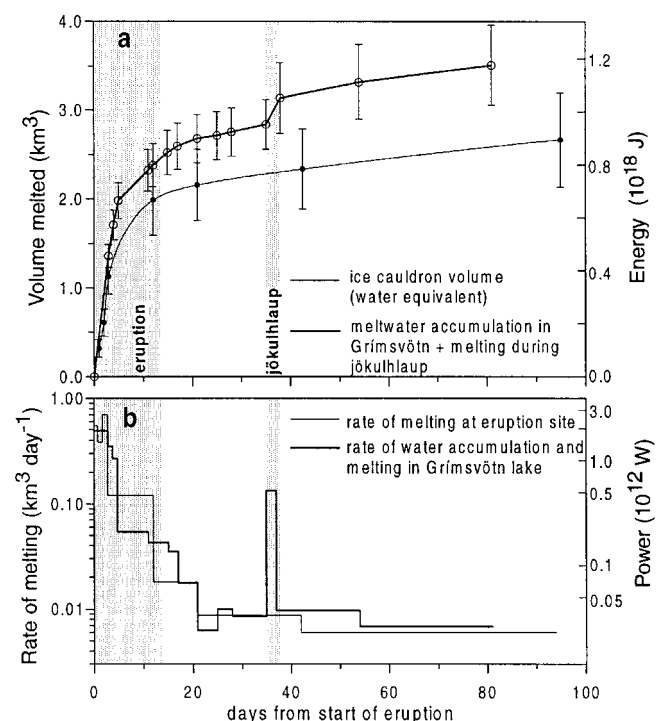
$$Q_e = M_{\text{ice}}L_{\text{ice}} + M_wC_wT_w \quad (2)$$

Figure 3 Ice melting and heat transfer rate from 1 October 1996 until 3 January 1997. **a**, Melting by the eruption could be measured by two independent methods: first, from the water equivalent volume of ice cauldrons above the eruptive fissure (obtained by radar altimetric measurements from an aeroplane equipped with a navigation unit); second, from the volume of meltwater in the Grímsvötn lake (by measuring the elevation of the lake's ice shelf with differential Global Positioning System measurements (DGPS) and knowledge of bedrock topography). Volume of the ice cauldrons is lower because it does not include the volume of the volcanic pile nor that of water melted at Grímsvötn by heat of the meltwater from the eruption site. To obtain total melting by the eruption, the volume of water stored in the porous volcanic pile, estimated as 0.3 km^3 , should be added to the accumulation curve for Grímsvötn. During the eruption 3.0 km^3 of ice was melted and in the following three months, 0.8 km^3 melted because of cooling of erupted material and the release of heat as pyroclastic glass was altered into consolidated hyaloclastite. During the jökulhlaup on 4–7 November a further 0.4 km^3 was melted by the remaining heat of the water, mainly at the subglacial outlet from Grímsvötn. The level of Grímsvötn fell by 175 m in the jökulhlaup when a total of 3.5 km^3 of water was drained from the lake. **b**, Rate of ice melting in water equivalent volume based on the same data as in **a**. In order to obtain the rate of melting at the eruption site from the ice cauldron data, it is assumed that volume of the volcanic pile increased from 0 to 0.4 km^3 in the first three days and from 0.4 to 0.7 km^3 in the next 10 days.

where M_{ice} and M_w are the masses of ice and meltwater respectively, T_w is the temperature of the meltwater, $L_{\text{ice}} = 335\text{ kJ kg}^{-1}$ is the latent heat of ice melting, and $C_w = 4.2\text{ kJ kg}^{-1}\text{ K}^{-1}$ is the heat capacity of water near 0°C . The amount of erupted magma can be estimated from equations (1) and (2) if the volume of melted ice is known. The volume of ice melted during the eruption and the following six weeks was 4.0 km^3 , requiring $1.1 \times 10^{12}\text{ kg}$ of magma if $\Delta T_e = 1,000^\circ\text{C}$, $T_w = 0^\circ\text{C}$ and efficiency was 100%. This equals 0.4 km^3 of magma, making the 1996 Vatnajökull eruption the fourth largest in Iceland this century. The dry density of the pile of volcanic glass is likely to be $1,500\text{--}1,600\text{ kg m}^{-3}$, similar to that of the Surtsey hyaloclastite²¹. The volume of the eruptive products under the ice cap is then $0.7\text{--}0.75\text{ km}^3$. In June 1997 the top of the new hyaloclastite mountain was observed within the ice canyon at the site of the subaerial crater in October. The exposed part was a 300-m-long ridge made of partly altered pyroclastic glass. Isolated radio-echo point soundings made in June 1997 suggest that the volcanic edifice is a 6–7-km-long ridge, similar to that shown in Fig. 2.

Highly turbulent convection of a mixture of meltwater and quenched ash fragments is called for to obtain the observed high rate of heat transfer. This mixture had a mean temperature of at least $15\text{--}20^\circ\text{C}$, the temperature required to account for the volume of ice melted over the subglacial path of the meltwater. Such heat in the meltwater greatly increases the rate of enlargement of subglacial tunnels²² and must be a key factor in the swiftness and destructive power of volcanic jökulhlaups. The meltwater drained towards Grímsvötn, apparently at the same rate as it was formed (Fig. 3a), precluding any significant water accumulation at the eruption site. Therefore the eruptive products mostly piled up against ice walls, containing the emerging volcano, and the walls remained steep as melting was compensated for by rapid ice flow towards the eruptive site. Thus, hyaloclastite mountains formed within ice caps may be steeper and not as laterally extensive as hyaloclastite mountains or islands of similar volume formed in water.

The product of the Vatnajökull eruption is a subglacial hyaloclastite ridge, partly located on top of a ridge formed in a similar eruption in 1938 (refs 22, 23; Fig. 2), increasing its maximum relief from 200 m to 500 m. Figure 4 shows schematically how the



eruption penetrated the ice cap by melting a narrow chimney through the ice. The subaerial eruption was fed through this chimney in the ice, 200–300 m wide and 50–100 m high, for 12 days without significant widening of the chimney. This gives insight into the conditions needed to form table mountains within ice caps. Lava caps on table mountains in Iceland^{6–8} and British Columbia⁵ indicate that openings melted in the overlying ice by eruptions must have had diameters of a few kilometres. This suggests that only prolonged or repeated eruptions at the same place can create the conditions required for table mountain formation within glaciers. Initial eruption of pillow lavas may also be important. Gradual heat release from a pile of pillow lava may lead to slower but more widespread melting, causing widening of any chimney formed through the ice cap in early stages of an eruption. Further volcanic activity at the 1996 eruption site might lead to the formation of a table mountain that would persist as a nunatak in the glacier.

It has been suggested that subglacial volcanism may play a role in the dynamics of West Antarctic ice streams by supplying water to their base²⁴. At Vatnajökull no rapid basal sliding on a regional scale was associated with the eruption. The formation of ice cauldrons over the eruptive fissure, and their subsequent widening, is the ice cap's response to sudden removal of mass from its base. Intense surface crevassing is similar to that observed in glacier surges but ice flow is probably dominated by internal deformation, not basal

sliding as in surges²⁵. The crevassed subsidence structure at the eruption site widened from 4 to 7 km in the three months after the eruption, reaching a width of 8 km in June 1997. Otherwise the surface of Vatnajökull remained intact with no signs of increased ice flow such as subsidence or crevassing. The effects of meltwater are also localized: it drains away either continuously, as at the 1996 eruption site, or in a sudden jökulhlaup as observed when the meltwater escaped from the Grímsvötn lake on 4–7 November. The removal of meltwater coupled with the formation of deep depressions may create steep gradients in basal water pressure, towards the depressions, and decrease water pressure at the base of the glacier surrounding the eruption site. This would impede basal sliding. Moreover, the mass removal by melting and drainage will decrease ice flow along the flowline downstream from the eruption site, as the glacier will act to heal itself by ice flow into the depression. Overall, our data suggest that at least for warm-based glaciers, the effects of subglacial volcanic eruptions are localized, with eruptions forming deep depressions and causing jökulhlaups. Significant changes in extent and shape of an ice sheet would require extensive, voluminous subglacial volcanism, melting a considerable fraction of the total ice volume in a short period of time. □

Received 14 April; accepted 29 August 1997.

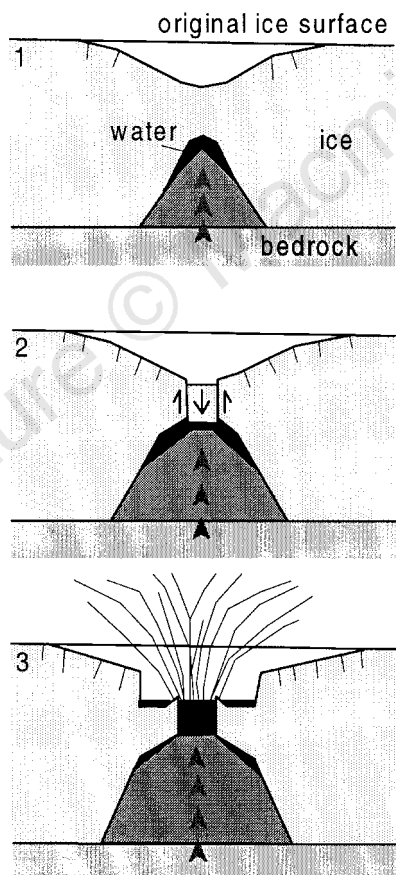


Figure 4 Schematic sections showing how the eruption penetrated the ice cap, through a narrow chimney through the ice. As the eruption melted its way through the glacier and the volcanic pile grew higher, the thickness of the overlying ice decreased until a block, 200–300 m in diameter, collapsed onto the underlying erupting vent. When this piston of ice had melted completely, the subaerial eruption commenced. Any melting of the chimney walls was apparently compensated by ice flow, as the width of the chimney remained 200–300 m despite continuous eruption for almost two weeks.

- Major, J. J. & Newhall, C. H. Snow and ice perturbation during historical volcanic eruptions and the formation of lahars and floods. *Bull. Volc.* **15**, 1–27 (1989).
- Lowe, D. R. *et al.* Lahars initiated by the 13 November 1985 eruption of Nevado del Ruiz, Colombia. *Nature* **324**, 51–53 (1986).
- Thorarinnsson, S. *Vötnin stríð* (Menningarsjóður, Reykjavík, 1974).
- LeMasurier, W. E. in *Volcanoes of the Antarctic Plate and Southern Oceans* (eds LeMasurier, W. E. & Thomson, J. W.) 1–17 (AGU, Antarctic Research Series 48, 1990).
- Mathews, W. H. "Tuyas", flat-topped volcanoes in northern British Columbia. *Am. J. Sci.* **245**, 560–570 (1947).
- Kjartansson, G. *Árnesingasaga* (Árnesingafélagið, Reykjavík, 1943).
- van Bemmelen, R. W. *et al.* *Tablemountains of Northern Iceland* (Brill, Leiden, 1955).
- Jones, J. G. Intraglacial volcanoes of the Laugarvatn region, south-west Iceland—I. *Q. J. Geol. Soc. Lond.* **124**, 197–211 (1969).
- Smellie, J. L. & Skilling, I. P. Products of subglacial volcanic eruptions under different ice thicknesses: two examples from Antarctica. *Sedim. Geol.* **91**, 115–129 (1994).
- Werner, R., Schminke, H. U. & Sigvaldason, G. A new model for the evolution of table mountains: volcanological and petrological evidence from Herdubreid and Herdubreidartögl volcanoes (Iceland). *Geol. Rundsch.* **85**, 390–397 (1996).
- Jakobsson, S. P. On the consolidation and palagonitization of the tephra of the Surtsey volcanic island, Iceland. *Surtsey Research Progress Rep.* **VI**, 121–128 (Surtsey Research Society, Reykjavík, 1972).
- Jakobsson, S. P. Environmental factors controlling the palagonitization of the Surtsey tephra, Iceland. *Bull. Geol. Soc. Denmark* **27**, 91–105 (1978).
- Jakobsson, S. P. & Moore, J. G. Hydrothermal minerals and alteration rates at Surtsey volcano, Iceland. *Geol. Soc. Am. Bull.* **97**, 648–659 (1986).
- Einarsson, P. *et al.* Center of the Iceland hotspot experiences volcanic unrest. *Eos* **78**, 374–375 (1997).
- Björnsson, H. Subglacial water reservoirs, jökulhlaups and volcanic eruptions. *Jökull* **25**, 1–14 (1975).
- Yokoyama, I. Energetics in active volcanos. *Tokyo Univ. Earthq. Res. Inst. Bull.* **35**, 75–97 (1957).
- Shimozuru, D. Discussion on the energy partition of volcanic eruption. *Bull. Volcanol.* **32**, 383–394 (1968).
- Charmichael, I. S. E., Turner, F. J. & Verhogen, J. *Igneous Petrology* 128–130 (McGraw-Hill, New York, 1974).
- Moore, J. G. Mechanism of formation of pillow lava. *Am. Sci.* **63**, 269–277 (1975).
- Bacon, C. R. High temperature heat content and heat capacity of silicate glasses: Experimental determination and a model for calculation. *Am. J. Sci.* **277**, 109–135 (1977).
- Oddson, B. Rock quality designation and drilling rate correlated with lithology and degree of alteration in volcanic rocks from the 1979 Surtsey drill hole. *Surtsey Research Progress Rep.* **IX**, 94–97 (Surtsey Research Society, Reykjavík, 1982).
- Björnsson, H. Hydrology of ice caps in volcanic regions. *Soc. Sci. Islandica* **45**, (Visindafélag Íslendinga, Reykjavík, 1988).
- Gudmundsson, M. T. & Björnsson, H. Eruptions in Grímsvötn, Vatnajökull, Iceland, 1934–1991. *Jökull* **41**, 21–45 (1991).
- Blankenship, D. D. *et al.* Active volcanism beneath the West Antarctic ice sheet and implications for ice-sheet stability. *Nature* **361**, 526–529 (1993).
- Kamb, B. *et al.* Glacier surge mechanism: 1982–1983 surge of Variegated Glacier, Alaska. *Science* **227**, 469–479 (1985).
- Björnsson, H. & Einarsson, P. Volcanoes beneath Vatnajökull, Iceland: Evidence from radio echo-sounding, earthquakes and jökulhlaups. *Jökull* **40**, 147–168 (1990).
- Björnsson, H., Björnsson, S. & Sigurgeirsson, Th. Penetration of water into hot rock boundaries of magma at Grímsvötn. *Nature* **295**, 580–581 (1982).
- Gudmundsson, M. T., Björnsson, H. & Pálsson, F. Changes in jökulhlaup sizes in Grímsvötn, Vatnajökull, Iceland, 1934–91, deduced from in-situ measurements of subglacial lake volume. *J. Glaciol.* **41**, 263–272 (1995).

Acknowledgements. We thank the Iceland Civil Aviation Authority, especially chief pilot S. Gudbjörnsson, the crews of the helicopters of the Icelandic Coast Guard and J. K. Björnsson of the Iceland Helicopter Service for enthusiastic cooperation during the monitoring of the eruption. F. Pálsson, H. Halldórsson, K. Pálsson, S. Jónsson and P. Simonarson took part in the GPS measurements and members of the Iceland Glaciological Society assisted in fieldwork in June 1997. The figures were prepared by Th. Högnadóttir. A review by D. Blankenship improved the paper.

Correspondence and requests for materials to M.T.G. (e-mail: mtg@raunvis.hi.is).

Extraordinary lifespans in ants: a test of evolutionary theories of ageing

Laurent Keller*† & Michel Genoud*

* Institute of Zoology and Animal Ecology, University of Lausanne, Bâtiment de Biologie, 1015 Lausanne, Switzerland

† Institute of Zoology, University of Bern, Ethologische Station Hasli, 3032 Hinterkappelen, Switzerland

Senescence presents not only a medical problem, but also an evolutionary paradox because it should be opposed by natural selection. Evolutionary hypotheses propose that ageing evolves as the necessary cost of processes increasing early reproductive success^{1,2}, or because of weaker selection against late-acting mutations³. A prediction of these hypotheses is that the rate of ageing should increase and the average lifespan decrease as the rate of extrinsic mortality increases^{1–7}. Alternatively, non-adaptive, purely mechanistic hypotheses invoke damage to DNA, cells, tissues and organs as being the unique cause of senescence and ineluctable death of organisms⁸. Here we show that the evolution of eusociality is associated with a 100-fold increase in insect lifespan. Such an increase is predicted by evolutionary theories because termite, bee and ant queens live in colonies that are sheltered and heavily defended against predators. Moreover, a comparison of ants with contrasting life histories also reveals an association between lifespan and extrinsic rate of mortality. These results provide strong support for evolutionary theories of ageing, as purely mechanistic hypotheses of senescence do not propose any association between the rate of extrinsic mortality and lifespans.

The physiological deterioration of organisms as they age presents an evolutionary paradox: if organisms can function well in youth, why can they not continue to do so in old age⁹? Evolutionary and non-adaptive, purely mechanistic hypotheses have been developed separately to account for the apparently universal occurrence of ageing in vertebrates and arthropods² and, with the exception of some experimental studies conducted with several *Drosophila* species, few attempts have been made to distinguish them^{2,9}. Evolutionary hypotheses predict that ageing rate should generally increase and average lifespan should decrease as the rate of age- and condition-independent (extrinsic) mortality increases and that it should be higher in those organisms whose fecundity does not increase markedly after maturity compared to those showing such an increase^{1,3–6,10}. In contrast, and all else being equal, purely physiological theories of senescence predict no relationship between extrinsic mortality, age-dependent fecundity, and rate of senescence. Instead, they frequently predict an association between duration of life and metabolic rate, whereby organisms with lower metabolic rates should have greater lifespans^{8,11}. It is important to note that evolutionary hypotheses do not deny that proximate mechanisms such as DNA damage are responsible for ageing: they only predict that natural selection will act on these proximate mechanisms. For example, a lower rate of extrinsic mortality should decrease the frequency of late-acting genes that reduce the efficiency of DNA repair and cause greater DNA damage.

Owing to their peculiar life history, highly eusocial insects with specialized worker and queen morphological castes (ants, termites and the honeybee) provide a unique opportunity to disentangle evolutionary and mechanistic hypotheses for ageing. Mortality of adult queens due to external factors such as predation is low because the queen is sheltered in a heavily defended nest. Hence, the queen is usually immune to predators beyond the nest-founding stage¹²,

except in species with unstable nests (see below). Moreover, because colonies typically start producing reproductives (males and new queens) only when the colony has reached a given size, queens experience a dramatic age-dependent increase in reproductive success when they initiate a new colony independently (without the help of workers). In such independent founding species, queens generally produce the first fertile offspring only when they are 3 to 4 years old¹², thus delaying the onset of ageing. Furthermore, the number of reproductives produced increases with colony size, which in turn correlates with the queen age¹². Evolutionary theories thus predict that queens of highly eusocial species such as ants, termites and honeybees should exhibit low rates of ageing and long lifespans. A further advantage of studying ageing in eusocial insects is that contrasting life histories often occur among species of similar body mass. Body mass has been shown to be correlated with lifespan and it is therefore a potentially confounding variable in comparative analyses of longevity^{6,13}.

As predicted by evolutionary theories, the lifespan of adult eusocial insect queens is much longer than that of adult solitary insects: the mean average longevity of ant, termite and honeybee queens is 10.0 ± 6.6 ($n = 51$), 11.5 ± 5.4 ($n = 9$), and 5.6 years, respectively (overall mean for eusocial insects 10.1 ± 6.4 years; $n = 61$), whereas that of 87 solitary insect species from eight orders is only 0.1 ± 0.2 years (Fig. 1). To remove the problem of phylogenetic non-independence between species^{14,15}, we compared the lifespan of eusocial insects with that of other insects by using an independent contrasts approach^{14,16}. This analysis reveals a significant correlation between longevity and eusociality ($P = 0.02$) among the 148 species for which data were available (Fig. 2). Purely mechanistic hypotheses cannot satisfactorily explain this difference. Ant, termite and honeybee queens are unlikely to have lower metabolic rates than non-eusocial insects which could account for such a difference: queens produce hundreds or even thousands of eggs every day which may represent more than their own body weight. This enormous reproductive effort must be associated with a high metabolic rate in the queen¹².

Additional support for the evolutionary hypotheses comes from the comparison among ant species with alternative reproductive strategies. Ant colonies may contain one (monogyny) or several queens (polygyny)¹⁷. Interspecific variation in queen number is generally associated with differences in nest-founding strategies. Queens of monogynous species typically initiate new colonies independently (that is, without the help of workers), whereas queens of polygynous species generally return or remain in an established colony after mating^{18–21}. The high risk of a colony losing

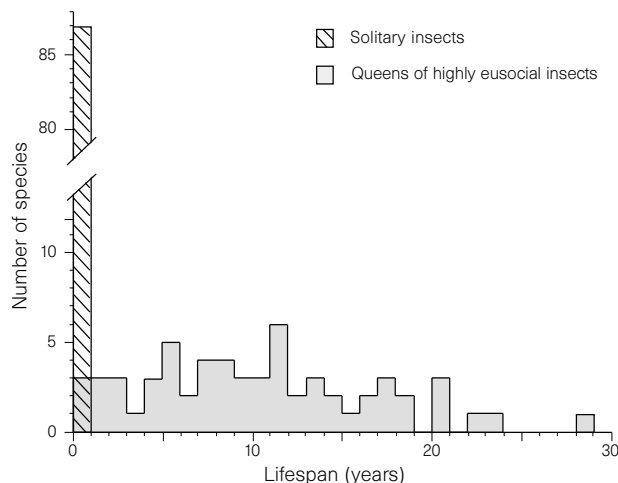


Figure 1 Mean lifespans of 87 solitary insect species from eight orders and 61 species of eusocial insects. We found data on queen lifespan in 51 ant species (Fig. 3), nine termite species and the honeybee.

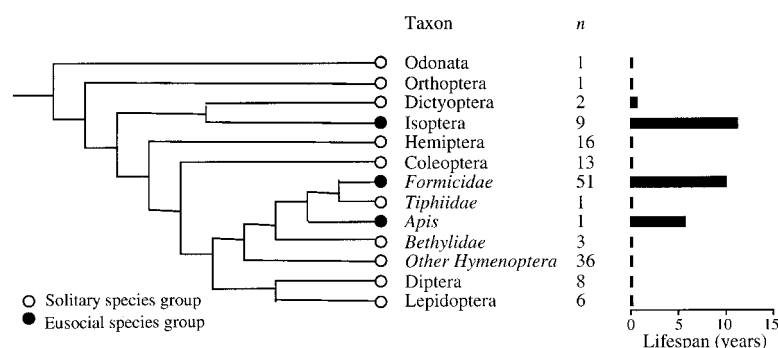


Figure 2 Average lifespans of adults from eight solitary orders of insects and queens from three eusocial insect taxa. Details on the order Hymenoptera is given for several families because eusociality evolved independently in the genus *Apis* and in Formicidae (ants). The complete phylogeny used for the comparative analysis of independent contrasts is available from L.K.

a queen is one of the major selective forces favouring polygyny over monogyny^{22,23}. For instance, polygyny is common in species that do not produce complex sheltered nests and frequently change nest location in response to environmental variation^{12,20}. Thus, in polygynous species the risk of extrinsic rate of mortality in mature colonies should be higher than in monogynous species. Moreover, queens returning to established nests after mating may produce sexuals sooner than queens that found independently¹⁸, hence their reproductive success is less age-dependent (see above).

Evolutionary hypotheses thus predict that a higher lifespan should be associated with monogyny and independent colony founding than with polygyny and dependent colony founding. In accord with this prediction, the average lifespan of queens is much greater in monogynous (12.3 ± 5.5 years; $n = 37$) than polygynous (1.6 ± 1.8 years; $n = 6$) species. The lifespan of queens in species with both monogynous and polygynous colonies is intermediate (5.7 ± 6.6 years; $n = 8$). Queen lifespan is also much greater in independent (12.2 ± 5.5 years; $n = 38$) than dependent (1.7 ± 1.6 years; $n = 7$) founding species, those species with mixed strategies exhibiting again intermediate values (5.4 ± 7.2 years; $n = 6$). Evolutionary changes in lifespan appear to have accompanied changes in queen number and mode of colony founding on several independent occasions (Fig. 3). An analysis correcting for the effect of phylogeny reveals that both queen number ($P = 0.001$) and mode of colony founding ($P = 0.0004$) are significantly correlated with lifespan. Because variation in queen number is tightly correlated with the mode of colony founding (Fig. 2), it is impossible to disentangle the relative effects of these two factors. However, this is not a problem because the evolutionary hypotheses predict that both monogyny and independent colony founding should select for increased lifespan, the pattern that our analysis revealed. Again, purely mechanistic hypotheses fail to explain this pattern. Although polygynous species change nest location much more frequently than monogynous species, such changes occur only a few times every year, which should not significantly affect the metabolic rate of queens. On the other hand, monogynous queens are usually much more fecund than their polygynous counterparts^{17–20}, thus possibly having an even higher metabolic rate.

Circumstantial evidence that longer maximum lifespan occurs in organisms that have lower mortality rates in the wild comes from the comparison of distantly related taxa. Birds and bats, which evade terrestrial predators by flight, generally have greater potential longevities than similarly sized terrestrial vertebrates (refs. 1, 2, but see ref. 24). Similarly, animals with thick shells, such as turtles and bivalves, tend to live longer than animals without armour^{1,2}. High maximum lifespans are also found in organisms whose fertility increases with age, such as fish and trees, as predicted by evolutionary theory². However, as already pointed out^{2,5,6}, these wide comparisons have shortcomings and should be considered as consistent with evolutionary theories of ageing rather than as rigorous tests. There are also experimental data supporting evolutionary theories of ageing, although their interpretation has been disputed. In experiments where lineages of *Drosophila* were selected

for late reproduction, longevity was increased^{22,25}. This response, which is predicted by evolutionary theories of ageing, is not easily explained by mechanistic theories.

In conclusion, our results offer a comparative analysis of lifespan data for eusocial insects which provides strong support for evolutionary but not purely mechanistic theories of ageing. The decrease in mortality rate and age-dependent increase in fecundity that has accompanied the evolution of eusocial life in ants and termites has

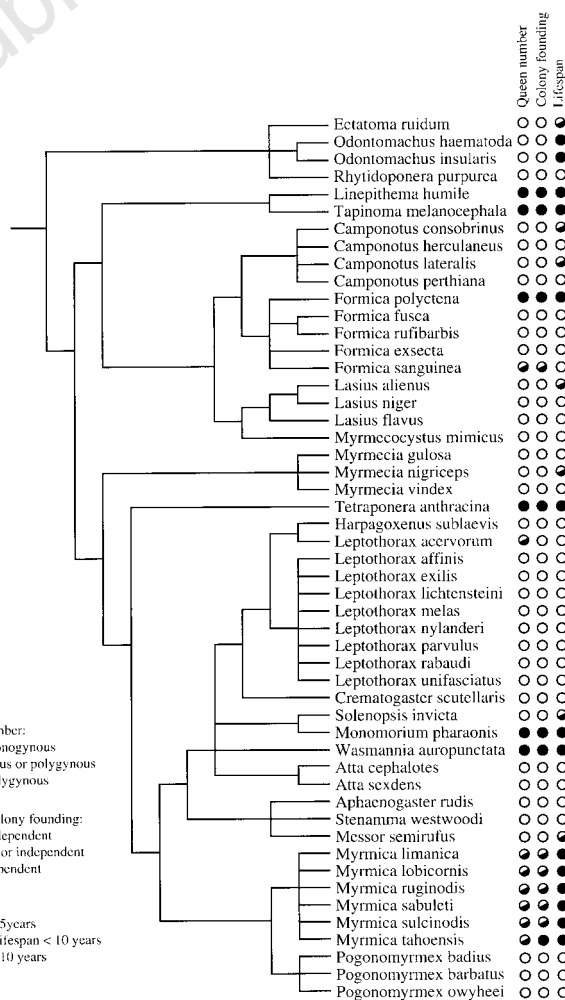


Figure 3 Average lifespans of ant queens in relation to their social organization. Species were classified as monogynous when nests always contain a single queen and as polygynous when nests contain several queens. Species in which both monogynous and polygynous nests occur were classified in an intermediate (mixed) category. Similarly, we differentiated three types of species depending on whether colony founding is always independent, always dependent or whether a mixture of both strategies occurs. References on the mode of colony founding and queen number per colony are available from L.K.

resulted in an over-100-fold increase in average maximum lifespan, with astonishing values of over 28 years for some ant species. Interestingly, the extrinsic mortality of the only known eusocial mammal, the naked mole-rat, is also low as a result of its subterranean life, and accordingly, the queens of this species live longer than any other rodent of comparable size²⁶. Finally, the comparison between ants with alternative life histories supports the assertion that differences in mortality rates and age-dependent fecundities, rather than any other factor associated with the evolution of social life, is responsible for these extreme lifespans. □

Methods

Comparative analyses. We used the CAIC (comparative analysis of independent contrasts) program, version 2.0.0 (ref. 16), which is based on the independent contrasts approach first proposed by Felsenstein¹⁴. Because our present knowledge of insect diversity is far from being exhaustive, only species used in our analysis were used to construct the phylogeny. Branch lengths were considered equal, which corresponds to a punctuational view of evolutionary change. Under this assumption, the CAIC performs acceptably well even when the phylogeny is poorly resolved²⁷. Standardization of the contrasts was checked according to the recommendations in ref. 28. All variables (independent and dependent) were considered continuous. For comparison among insects with different levels of eusociality, 'eusociality' took the value 1 for solitary insects and 2 for eusocial species. For the comparison among ants with contrasting life histories, values ranging between 1 and 3 were assigned to both queen number (monogynous, mixed or polygynous) and mode of colony founding (independent, mixed or dependent).

Lifespan data. In three ant and two termite species, only maximum lifespans of queens were given. For these species, we estimated mean lifespan by dividing maximum lifespan by 1.45, the ratio between maximum and average lifespan observed in *Aphaenogaster rudis* and *Myrmecia vindex* (the two species with the largest data set). In some cases the average lifespan is a minimum value as the age of the queen was unknown when the observations began and as queens were still alive when the data were published. Queens were apparently still reproducing just before their death. Data on the lifespan of non-eusocial insect species come from Ridley^{29,30}, the only large data set on insect lifespan. These data come from laboratory studies with almost natural rearing conditions. Most of the data on queen lifespan come from laboratory studies, but a minority of the data are from field studies (in 3 genera). Because extrinsic mortality may have occurred, these studies may yield an underestimate of average lifespan. This makes the comparison between eusocial and non-eusocial species conservative. For comparison between ants with contrasting life histories, we also conducted the comparative analysis of independent contrasts without considering species in which lifespan was estimated in the field. The exclusion of these species did not affect the significance of the tests performed. Exact lifespans of each species and references are available from L.K.

Received 23 May; accepted 25 July 1997.

- Williams, G. C. Pleiotropy, natural selection, and the evolution of senescence. *Evolution* **11**, 398–441 (1957).
- Rose, M. *Evolutionary Biology of Aging* (Oxford Univ. Press, Oxford, 1991).
- Hamilton, W. D. The moulding of senescence by natural selection. *J. Theor. Biol.* **12**, 12–45 (1966).
- Stearns, S. C. *The Evolution of Life Histories* (Oxford Univ. Press, Oxford, 1992).
- Partridge, L. & Barton, N. H. Optimality, mutation and the evolution of ageing. *Nature* **362**, 305–311 (1993).
- Partridge, L. & Barton, N. H. On measuring the rate of ageing. *Proc. R. Soc. Lond. B* **263**, 1365–1371 (1996).
- Charlesworth, B. *Evolution in Age-structured Populations* 2nd edn (Cambridge Univ. Press, Cambridge, 1994).
- Finch, C. E. *Longevity, Senescence and the Genome* (Chicago Univ. Press, Chicago, 1990).
- Bell, G. Evolutionary and nonevolutionary theories of senescence. *Am. Nat.* **124**, 600–603 (1984).
- Abrams, P. A. Does increased mortality favor the evolution of more rapid senescence? *Evolution* **47**, 877–887 (1993).
- Sohal, R. S. in *Insect Ageing* (eds Collatz, K.-G. & Sohal, R. S.) 23–44 (Springer, Berlin, 1986).
- Hölldobler, B. & Wilson, E. O. *The Ants* (Springer, Berlin, 1990).
- Promislow, D. E. L. Senescence in natural populations of mammals: a comparative study. *Evolution* **45**, 1869–1887 (1991).
- Felsenstein, J. Phylogenies and the comparative method. *Am. Nat.* **125**, 1–15 (1985).
- Harvey, P. H. & Pagel, M. D. *The Comparative Method in Evolutionary Biology* (Oxford Univ. Press, Oxford, 1991).
- Purvis, A. & Rambaut, A. Comparative analysis by independent contrast (CAIC): an Apple Macintosh application for analysing comparative data. *Comp. Appl. Biosci.* **11**, 247–251 (1995).
- Keller, L. (ed.). *Queen Number and Sociality in Insects* (Oxford Univ. Press, Oxford, 1993).
- Keller, L. The assessment of reproductive success of queens in ants and other social insects. *Oikos* **67**, 177–180 (1993).
- Keller, L. Social life: the paradox of multiple-queen colonies. *Trends Ecol. Evol.* **10**, 355–360 (1995).
- Bourke, A. F. G. & Franks, N. R. *Social Evolution in Ants* (Princeton Univ. Press, Princeton, 1995).

- Ross, K. G. & Keller, L. Ecology and evolution of social organization—insights from fire ants and other highly eusocial insects. *Ann. Rev. Ecol. Syst.* **26**, 631–656 (1995).
- Nonacs, P. Queen number in colonies of social Hymenoptera as a kin-selected adaptation. *Evolution* **42**, 566–580 (1988).
- Reeve, H. K. & Nonacs, P. Within-group aggression and the value of group members—theory and a field test with social wasps. *Behav. Ecol.* **8**, 75–82 (1997).
- Promislow, D. E. L. Senescence in natural populations of mammals: a comparative study. *Evolution* **45**, 1869–1887 (1991).
- Partridge, L. & Fowler, K. Direct and correlated responses to selection on age at reproduction in *Drosophila melanogaster*. *Evolution* **46**, 76–91 (1992).
- Jarvis, J. U. M. & Bennett, N. C. in *The Biology of the Naked Mole-rat* (eds Sherman, P. W., Jarvis, J. U. M. & Alexander, R. D.) 66–69 (Princeton Univ. Press, Princeton, 1991).
- Purvis, A., Gittleman, J. L. & Luh, H.-K. Truth or consequences: effects of phylogenetic accuracy on two comparative methods. *J. Theor. Biol.* **167**, 293–300 (1994).
- Garland, T. J., Harvey, P. H. & Ives, A. R. Procedures for the analysis of comparative data using independent contrasts. *Syst. Biol.* **41**, 18–32 (1992).
- Ridley, M. Mating frequency and fecundity in insects. *Biol. Rev.* **63**, 509–549 (1988).
- Ridley, M. Clutch size and mating frequency in parasitic Hymenoptera. *Am. Nat.* **142**, 893–910 (1993).

Acknowledgements. We thank P. Abrams, A. Bourke, W. Brown, M. Chapuisat, B. Charlesworth, M. Milinski, L. Partridge, N. Perrin, F. Ratnieks, M. Rose, K. Reeve, M. Ridley and S. Stearns for comments on the manuscript, and S. Cover for constructing the ant phylogeny. This work was funded by grants from the Swiss NSF (L.K.).

Correspondence and requests for materials should be addressed to L.K. (e-mail: Laurent.keller@izea.unil.ch).

The lateral line can mediate rheotaxis in fish

John C. Montgomery, Cindy F. Baker & Alexander G. Carton

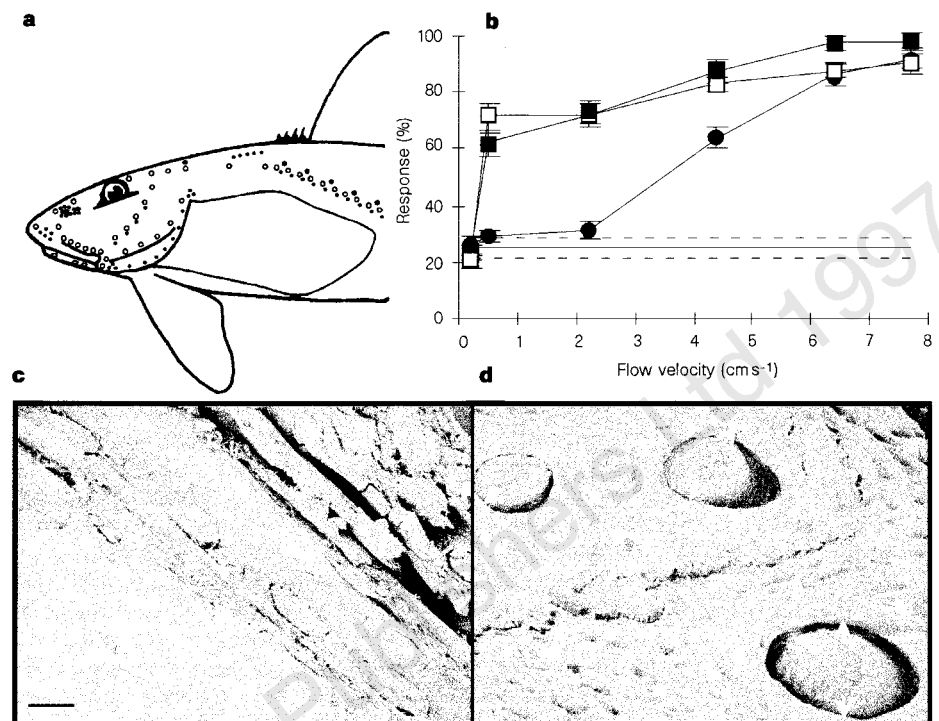
Experimental Biology Research Group, School of Biological Sciences, University of Auckland, Private Bag 92019 Auckland, New Zealand

Rheotaxis is a behavioural orientation to water currents¹. It has been demonstrated physiologically that some lateral-line receptors are particularly well suited to provide information on water currents², but their contribution to rheotaxis has been largely overlooked. The accepted view is that rheotaxis is mediated by visual and tactile cues¹, and that in rheotactic orientation "the lateral lines play only a minor role"³. Here we provide a direct demonstration that rheotaxis can be mediated by the lateral line, and indeed by one specific receptor class of this system. In three diverse fish species, pharmacological block of the entire lateral-line system substantially increases the velocity threshold for rheotactic behaviour. The same effect is observed when only superficial neuromasts are ablated, whereas blockade of the other receptor class, canal neuromasts, has no such effect. Our results therefore demonstrate that superficial neuromasts make an important contribution to rheotactic behaviour in fish.

Rheotaxis is an important and widespread behaviour in fish⁴. Examples include upstream migration, which counteracts the passive drift of eggs and larvae, and station-holding by facing upstream, which puts the fish in the best position to intercept food and odours carried in the current.

But what is the sensory basis of rheotaxis? A fish displaced by a current can use visual and tactile cues to initiate a response. Indeed, detection of the current has to be made against some stationary reference point. However, fish should also be able to detect current flow using the mechanosensory lateral-line system. This system, which is found in all fish, has receptors that are hair-cell patches called neuromasts. There are two main receptor classes within the lateral line: canal neuromasts, which are located within canals under the skin, and superficial neuromasts, which occur on the surface of the skin. Canal neuromasts respond best to water accelerations², such as those produced by the movements of other animals, whereas electrophysiological recordings show that superficial neuromasts respond to the velocity of water movements over the skin of the fish⁵. However, it has not been demonstrated clearly that the information from superficial neuromasts has an effect on the

Figure 1 Lateral-line sense organs and rheotactic responses of the torrentfish. **a**, Lateral view of the torrentfish; small dots, superficial neuromasts; open circles, canal pores. **b**, Rheotactic responses; horizontal solid and broken lines indicate mean $\pm$ 95% confidence intervals of the orientation response in the absence of current; filled squares, normal fish; open squares, gentamicin-treated fish; filled circles, cobalt-treated fish. **c**, Scanning electron micrograph of the canal neuromasts showing a section of the canal opened longitudinally to expose the canal neuromasts. **d**, Scanning electron micrograph of superficial neuromasts from the area of the operculum; white arrowheads indicate the direction of orientation of the hair cells.



behaviour of the fish. The obvious candidate behaviour for study is rheotaxis. Here we show that the superficial neuromasts can indeed make a significant contribution to rheotactic orientation, particularly at low current velocities.

Experiments were conducted on three species of fish: the torrentfish *Cheimarrichthys fosteri*, which are stream-dwelling, nocturnally active lateral-line predators⁶; the antarctic fish *Pagothenia borchgrevinkii*, which use their lateral line to detect planktonic prey⁷; and the blind cave-fish *Astyanax fasciatus*, which have been used extensively in studies of lateral line-mediated hydrodynamic imaging⁸. As the cave-fish swim they produce a flow field which is distorted by stationary objects in the water around them, and these distortions are detected by the lateral line. All three species have both canal and superficial neuromast systems in the lateral line. The superficial neuromasts typically occur individually, or in rows, on the surface of the head and the trunk (Figs 1, 2), although in the cave-fish there is an extensive arrangement of superficial neuromasts over the surface of the head and body (Fig. 3). The neuromasts themselves consist of a patch of hair cells with overlying cupula (Fig. 1c, d), and displacement of the cupula by surrounding water movements forms the effective stimulus to the system. All three species have a series of fluid-filled subdermal canals over the surface of the head and along the trunk that open to the exterior through pores, with typically one neuromast between each pair of pores (Fig. 1c).

The fish were housed in flumes and rheotactic orientation was observed and recorded. Details (see Methods) of the rheotactic behaviour and measurement of the response were different across the three species. The overall pattern of the response was remarkably similar (Figs 1–3), all three species showing an unconditioned rheotactic response. The threshold for a rheotactic response in torrentfish, antarctic fish and cave-fish was <0.5 , <2 and <3 cm s⁻¹, respectively, but the threshold response for the blind cave-fish reduced to <0.4 cm s⁻¹ in the presence of a diffuse olfactory cue. Cobalt treatment (torrentfish and cave-fish) and streptomycin treatment (antarctic fish) was used to block the entire lateral line system^{9,10}, resulting in a marked elevation of the threshold for rheotactic behaviour. This effect was not due to blockade of the canal neuromast system because the original rheotactic response was unaltered by treatment with gentamicin,

which has been shown to selectively disrupt canal neuromasts¹¹, an effect that was verified by scanning electron microscopy in the cave fish. Gentamicin-treated canal neuromasts were without apical cilia, whereas the superficial neuromasts appeared normal. In contrast, physical ablation of the superficial neuromasts in antarctic fish and cave-fish produced rheotactic behaviour that was identical to that produced by pharmacological block of the entire lateral line.

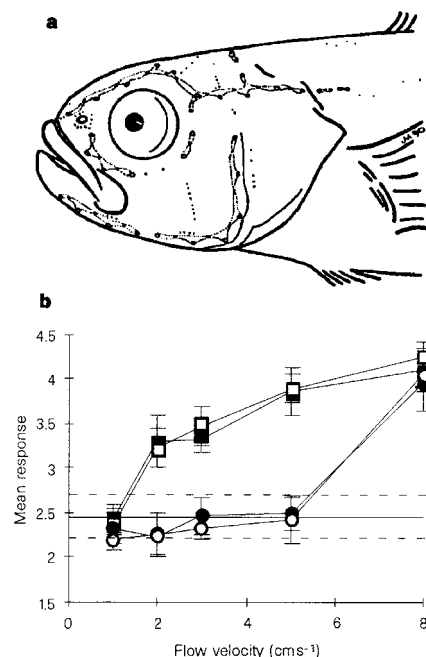


Figure 2 Lateral-line sense organs and rheotactic responses of the antarctic fish. **a**, Distribution of lateral-line endorgans¹⁵; small dots, superficial neuromasts; larger dots, canal pores. **b**, Rheotactic responses; horizontal solid and broken lines indicate mean $\pm$ 95% confidence intervals of the orientation response in the absence of current; filled squares, normal fish; open squares, gentamicin-treated fish; filled circles, streptomycin-treated fish; open circles, after physical ablation of the superficial neuromasts. Mean response, no. of fish facing upstream.

These results demonstrate unequivocally that the superficial neuromasts have a clear effect on rheotactic behaviour. It was almost entirely predictable from the known physiological responses of superficial neuromasts that the superficial neuromast system was involved in rheotactic behaviour, but it took this direct demonstration to challenge the historical view of the sensory basis of rheotaxis^{1,3}.

It remains true that orientation to current requires a fixed external reference point. For fish in the water column this is best provided visually, and vision may well dominate station-holding behaviour¹. However, even in this situation the lateral line may be providing information on water velocity and direction. The absence of a direct behavioural response does not necessarily imply that the information is not available to the fish to be used when required. For example, the reduction in rheotactic threshold for the cave-fish with the addition of an olfactory cue shows that they can detect slower water flows than those that typically elicit the unconditioned rheotaxis. This observation also provides evidence for a potential functional linkage between lateral lines and olfaction¹².

In the absence of vision, intermittent or continuous contact with the substrate can provide a tactile reference frame against which current can be detected. At higher current speeds, contact slippage can be a sufficient cue for rheotaxis. The important contribution of the superficial lateral-line system, in this instance, is to provide information at lower current velocities.

The three species studied were chosen to illustrate the contribution of lateral-line inputs to rheotaxis over a range of benthic to semipelagic species. Torrentfish are a nocturnal benthic species in which orientation to current facilitates the lateral-line detection of invertebrate stream drift. With this species experiments were con-

ducted under infrared light to exclude the possibility that the fish used visual cues. Visual cues are clearly not available to blind cave-fish, which are a semipelagic species dependent on intermittent substrate contact to provide an external frame of reference. The antarctic fish is generally described as cryopelagic, swimming in the area beneath the fast ice of the high-latitude antarctic sounds, although in the experimental flume they typically maintained contact with the substrate. Even though these experiments were conducted in the light, with visual cues available, this species exhibited the same pattern of lateral line-mediated rheotactic behaviour as the other two species.

Thus orientation to water currents, like other orientation capabilities of animals¹³, is mediated by a multiplicity of sensory cues. The cues available and their relative importance will vary with circumstance, but it makes sense that hydrodynamic receptors, such as the superficial neuromasts of the lateral line, should be important in rheotactic behaviour. □

Methods

Rheotactic measures. Different rheotactic measures were used owing to the different behaviour of the three species. For the torrentfish (40–52 mm long), rheotaxis was measured as the percentage of fish ($n = 5$) facing within 45° of the upstream direction, with 10 observations taken at 2-min intervals once the flow was established. For the antarctic fish *P. borchgrevinki* (190–210 mm long), rheotaxis was measured as the mean number of fish ($n = 5$) facing within 20° of the upstream direction, with 10 observations taken at 2-min intervals 10 min after the flow was initiated. For the cave-fish (40–70 mm long), responses were measured as the mean number of fish ($n = 5$) facing directly upstream with 10 observations taken at 30-s intervals starting 5 min after a new flow was initiated.

Cobalt treatment. Pharmacological blockade of lateral-line receptors using cobalt has been well established for use in freshwater species⁹. Fish were held for 3 h in a 2 mmol l⁻¹ cobalt solution made up in calcium-free water. This treatment was repeated at 4-day intervals if experiments extended over this period.

Streptomycin treatment. For the marine species *P. borchgrevinki* we used streptomycin treatment (0.5 g l⁻¹ for 3 h). The use of streptomycin to block lateral lines in marine species has been reviewed¹⁴. Unlike cobalt treatment, there is no direct demonstration that external application of streptomycin does not affect inner-ear receptors in fish. However, *P. borchgrevinki* sat on the bottom of the flume and exhibited rheotaxis at current levels well below those that produced whole-body movement. That is, the rheotactic behaviour was observed in the absence of any inner-ear stimulus, so the question of a confounding effect of streptomycin on inner-ear function does not arise.

Gentamicin treatment. Gentamicin has been shown to morphologically damage canal hair cells but leave hair cells of superficial neuromasts intact and this pattern of damage was confirmed in our study. Our protocol for gentamicin treatment was to immerse the fish in a 0.002% solution for 1–4 days as described¹¹. Direct experimental verification that superficial neuromasts remain functionally intact after gentamicin treatment was lacking, but our results show that superficial neuromasts remain fully functional. Otherwise, gentamicin treatment would have been expected to partly exhibit the same effect as physical ablation of the superficial neuromasts.

Ablation of neuromasts. Superficial neuromasts are delicate structures, and gentle scraping of the surface of the skin was an effective method of physical ablation. For the cave-fish, the extensive coverage of superficial neuromasts meant that two passes on successive days were required to completely disable the superficial neuromast system. Partial ablation produces a partial effect in cave-fish, with a rheotactic threshold intermediate between intact fish and those with complete blockade of the lateral line. Physical ablation of the neuromasts in antarctic fish is a much more specific procedure, owing to the limited distribution of superficial neuromasts in this species. Sham operations on the antarctic fish have no effect on rheotaxis, excluding the possibility that the effect is brought about by a general impairment to the tactile sense.

Behavioural analysis. Removal of specific sensory modalities could influence motivational state. However, any behavioural change always occurred directly after treatment, and continued for as long as was tested (10 days in the torrentfish). Responses were consistent across the three species and were

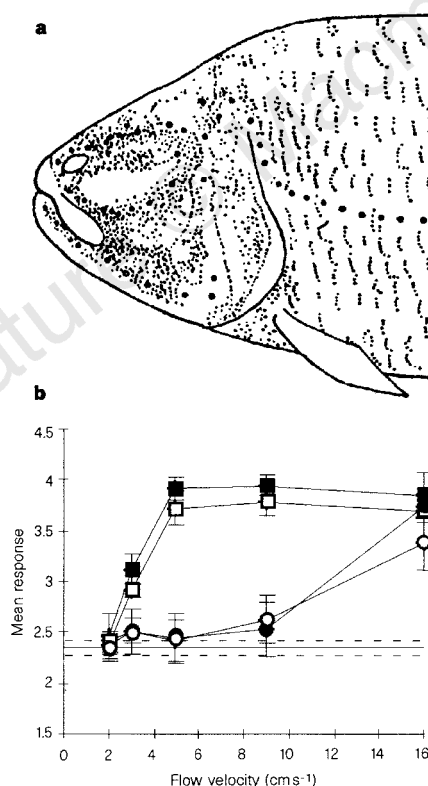


Figure 3 Lateral-line sense organs and rheotactic responses of the blind cave-fish. **a**, Distribution of lateral-line endorgans¹⁶; small dots, superficial neuromasts; larger dots, canal pores. Note the extensive covering of superficial neuromasts on both the head and the trunk. **b**, Rheotactic responses; horizontal solid and dotted lines indicate mean $\pm$ 95% confidence intervals of the orientation response in the absence of current; filled squares, normal fish; open squares, gentamicin-treated fish; filled circles, cobalt-treated fish; open circles, after physical ablation of the superficial neuromasts.

constant within each species. All fish continued to swim and feed visually with no evidence of a lack of motivation.

Received 29 February; accepted 1 August 1997.

1. Arnold, G. P. Rheotrophism in fishes. *Biol. Rev.* **49**, 515–576 (1974).
2. Münz, H. in *The Mechanosensory Lateral Line Neurobiology and Evolution* (eds Coombs, S., Gorner, P. & Münz, H.) 285–298 (Springer, New York, 1989).
3. Dijkgraaf, S. in *The Mechanosensory Lateral Line Neurobiology and Evolution* (eds Coombs, S., Gorner, P. & Münz, H.) 7–14 (Springer, New York, 1989).
4. Montgomery, J. C., Coombs, S. & Halstead, M. Biology of the mechanosensory lateral line in fishes. *Rev. Fish Biol. Fish.* **5**, 399–416 (1995).
5. Coombs, S. & Montgomery, J. C. Function and evolution of superficial neuromasts in an antarctic nototheniid fish. *Brain Behav. Evol.* **44**, 287–298 (1994).
6. Montgomery, J. C. & Milton, R. C. Use of the lateral line for feeding in the Torrent fish (*Cheimarrichthys fosteri*). *N. Z. J. Zool.* **20**, 121–125 (1993).
7. Montgomery, J. C. & Macdonald, J. A. Sensory tuning of lateral line receptors in Antarctic fish to the movements of planktonic prey. *Science* **235**, 195–196 (1987).
8. Abdel-Latif, H., Hassan, E. S. & von Campenhausen, C. Sensory performance of blind Mexican cave fish after destruction of the canal neuromasts. *Naturwissenschaften* **77**, 237–239 (1990).
9. Karlsen, H. E. & Sand, O. Selective and reversible blocking of the lateral line in freshwater fish. *J. Exp. Biol.* **133**, 249–262 (1987).
10. Kroese, A. B. A. & van den Bercken, J. Effects of ototoxic antibiotics on sensory hair cell functioning. *Hearing Res.* **6**, 183–197 (1982).
11. Song, J., Yan, H. Y. & Popper, A. N. Damage and recovery of hair cells in fish canal (but not superficial) neuromasts after gentamicin exposure. *Hear. Res.* **91**, 63–71 (1995).
12. Montgomery, J. C. & Pankhurst, N. W. in *Deep-sea Fish* (eds Randall, D. J. & Farrell, A. P.) 325–349 (Academic, San Diego, 1997).
13. Able, K. P. & Able, M. A. Interactions in the flexible orientation system of a migratory bird. *Nature* **375**, 230–232 (1995).
14. Baxter, J. H. S. & Fuiman, L. A. in *The Mechanosensory Lateral Line Neurobiology and Evolution* (eds Coombs, S., Gorner, P. & Münz, H.) 481–499 (Springer, New York, 1989).
15. Andriashev, A. P. & Jakubowski, M. Morphological grounds for generic separation of the Antarctic broadhead-fishes (*Trematomus borchgrevinkii*) Boulenger and *T. brachysoma* Pappenheim and a new status of the genus *Pagothenia* Nichols & Lamonte (Nototheniidae). *Zool. Zh.* **50**, 1041–1055 (1971).
16. Schemmel, C. Vergleichende Untersuchungen an den Hautsinnesorganen ober- und unterirdisch lebender *Astyranax*-Formen ein Beitrag zur Evolution der Cavernicolen. *Z. Morphol. Tiere* **61**, 255–316 (1967).

Acknowledgements. We thank the Marsden Fund for support, Kelly Tarlton's Underwater World for an Antarctic scholarship and Carol Diebel for at times acting in loco mentis.

Correspondence and requests for materials should be addressed to J.C.M. (e-mail: j.montgomery@auckland.ac.nz).

Mutation of an axonemal dynein affects left-right asymmetry in *inversus viscerum* mice

Dorothy M. Supp*, David P. Witte†, S. Steven Potter* & Martina Brueckner‡

Divisions of *Development Biology and †Pathology, The Children's Hospital Research Foundation, Cincinnati, Ohio 45229, USA

‡Department of Pediatrics/Cardiology, Yale School of Medicine, New Haven, Connecticut 06520, USA

The development of characteristic visceral asymmetries along the left–right (LR) axis in an initially bilaterally symmetrical embryo is an essential feature of vertebrate patterning. The allelic mouse mutations *inversus viscerum* (*iv*)^{1,2} and *legless* (*lgl*)^{3,4} produce LR inversion, or situs inversus, in half of live-born homozygotes. This suggests that the *iv* gene product drives correct LR determination, and in its absence this process is randomized². These mutations provide tools for studying the development of LR-handed asymmetry and provide mouse models of human lateralization defects. At the molecular level, the normally LR asymmetric expression patterns of *nodal*⁵ and *lefty*⁶ are randomized in *iv/iv* embryos, suggesting that *iv* functions early in the genetic hierarchy of LR specification. Here we report the positional cloning of an axonemal dynein heavy-chain gene, *left/right-dynein* (*lrd*), that is mutated in both *lgl* and *iv*. *lrd* is expressed in the node of the embryo at embryonic day 7.5, consistent with its having a role in LR development⁷. Our findings indicate that dynein, a microtubule-based motor, is involved in the determination of LR-handed asymmetry and provide insight into the early molecular mechanisms of this process.

In humans, random development of left–right asymmetry is a feature of Kartagener's syndrome, which is manifested by situs inversus, immotile cilia and the absence of ciliary dynein arms; this suggests that dynein function has a role in LR development^{8,9}. Dyneins are a family of minus-end-directed microtubule-based motors which are traditionally classified as either cytoplasmic or axonemal. Cytoplasmic dyneins have been implicated in vesicular transport, nuclear migration, and spindle orientation^{10,11}, whereas axonemal dyneins produce the motive forces that cause the sliding of adjacent microtubules in the axoneme^{10,11} and thus produce ciliary and flagellar movement. Dyneins function as large multi-subunit complexes containing up to three heavy chains, which include the force-producing motor domain, in addition to several intermediate and light chains. So far, 15 distinct dynein heavy-chain genes have been identified in vertebrates^{12–14}. Two of these (*Dnchc1* and *Dnahc13*) were mapped to the distal region of mouse chromosome 12, close to the previously identified map location of the *iv* and *lgl* mutations^{4,15}.

We isolated and analysed overlapping YAC (yeast artificial chromosome) clones spanning the *lgl* transgene insertion site in this region of chromosome 12, which indicated that the *lgl* mutation includes a deletion of at least 600 kilobases (kb) (Fig. 1a). Our inability to isolate non-chimaeric YACs at one end of the contig or to identify the remaining mouse DNA–transgene junction suggests that this deletion may extend to the end of the chromosome. To determine whether a dynein gene is involved in the *iv* and *lgl* mutations, we screened the *lgl* YAC contig at low stringency with a DNA probe for the highly conserved first P-loop nucleotide-binding domain found in both cytoplasmic and axonemal dynein. This strategy identified a 2.1-kb genomic *Bgl*II fragment which cross-hybridizes to the dynein P-loop probe and is deleted in *lgl* DNA (see Supplementary information). Sequence analysis of this fragment revealed that it is distinct from *Dnchc1* and *Dnahc13* and includes two exons orthologous to a rat dynein-like gene, *DLP11* (ref. 12). We cloned over 7 kb of coding sequence from this gene, including all four central P-loops, and sequence analysis reveals 67% identity and 81% similarity to the sea-urchin axonemal dynein- β heavy-chain gene, compared with only 31% identity and 54% similarity to the rat cytoplasmic dynein heavy-chain gene (Fig. 1b). In

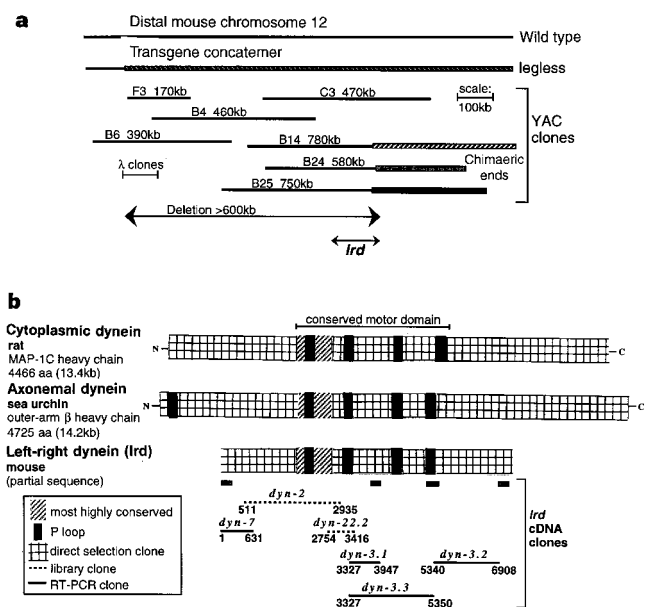


Figure 1 Identification of a dynein gene deleted in *lgl*. **a**, YAC contig spanning the mouse DNA/transgene junction in *lgl*. The location of the newly identified *lrd* gene (see text) is indicated. **b**, Homology of cloned *lrd* sequences with rat cytoplasmic and sea-urchin axonemal dynein heavy chains.

addition, the sequence immediately adjacent to the first P-loop (TETTKDL) defines this gene as an axonemal dynein isoform¹⁶. Recombination mapping using the same C57BL/6J X SI/Col-*iv* outcross/backcross panel initially used to determine the map position of *iv* (ref. 17) demonstrated no recombination between this gene, which we designate *left/right-dynein* (*lrd*), and *iv* out of 39 informative meioses, including the one animal that exhibited a recombination event between the distal end of IgH-V and *iv*¹⁷ (data not shown).

Comparison of *lrd* sequences from wild type and *iv* demonstrates a striking missense mutation in *iv* mice. Analysis of 7 kb of sequence obtained from cloned reverse transcription polymerase chain reaction (RT-PCR) products from wild-type (ICR outbred mice) and *iv*/*iv* homozygous mice revealed six silent base changes that did not alter coding potential, and are therefore probably strain polymorphisms. However, one difference was found that changes a glutamic acid to a lysine at a position corresponding to residue 2,249 of the sea-urchin sequence (Fig. 2a). This amino-acid substitution, resulting in a change from negative to positive charge, is located between the second and third P-loop motifs. This region of the protein is part of the highly conserved central portion of the heavy chain that constitutes the motor domain¹⁰ (Fig. 1b). The G → A base change in *iv* was confirmed by sequencing RT-PCR products from multiple RNA preparations and by sequencing the relevant exon from genomic DNAs of multiple *iv* and wild-type mice. The *iv* mutation arose in an outbred stock of mice that was subsequently crossed into the strains DBA, C3H and BALB/c; hence no parental strain is available for comparison. Thus, to establish that the mutation did not represent a strain polymorphism, we sequenced the genomic DNA encompassing this region of the *lrd* gene from eight different strains of laboratory mice (C57BL/6J, C3H, BALB/c, A/J, AKR/J, DBA, 129/J and SJL/J), the wild-derived strains TIRANO/Ei (*Mus domesticus*), SPRET/Ei (*Mus spretus*), and PANCEVO/Ei (*Mus hortulanus*), as well as rat (*Rattus norvegicus*). All contain the wild-type glutamic acid at this position. As shown in Fig. 2b, this region of the protein is highly conserved across widely divergent species and among both axonemal and cytoplasmic

dynein heavy chains. The glutamic acid residue at this position is absolutely conserved among all known dynein heavy-chain sequences, which suggests that it is critical to the function of the protein. We conclude that this glutamic acid-to-lysine substitution is unique to *iv* mutants and therefore provides a molecular basis for the randomization of LR development seen in these mice.

The *lgl* phenotype, with hindlimb, brain and craniofacial malformations, as well as situs inversus³, is more severe than the *iv* phenotype. This could be the result of a more severe *lrd* mutation in *lgl*—a deletion versus the missense mutation found in *iv*—and/or the result of additional genes mutated in *lgl* mice.

Expression of *lrd* messenger RNA was examined by northern blot hybridization, RT-PCR and *in situ* hybridization. Northern blot hybridization demonstrates a low abundance *lrd* transcript of ~15 kb in adult brain (see Supplementary information). Expression was detected at a lower level in adult testes (data not shown). The transcript size appears similar in *iv* and wild-type RNA. RT-PCR showed that *lrd* mRNA is expressed as early as the blastocyst stage (embryonic day (E) 3.5) and is also expressed in embryonic stem (ES) cells, which are derived from the blastocyst inner cell mass (Fig. 3). Expression is also seen at E8, E9 and E16.5 (Fig. 3). No transcripts were detected in mouse embryo fibroblasts (MEFs; Fig. 3). *In situ* hybridization showed that *lrd* mRNA embryonic expression is highly restricted; at E7.5, before the time of asymmetric expression of *nodal* and *lefty*, *lrd* mRNA is detected only in the ventral cells of the node (Fig. 4a–c). We detected no expression in the heart at any stage. These results show that the onset of *lrd* mRNA expression precedes that of *nodal* and *lefty*, two of the earliest molecular markers of LR asymmetry, which are misexpressed in *iv* mutant embryos.^{5,6} Also, the presence of *lrd* transcripts in the node implicates this mammalian equivalent of Spemann's organizer in the establishment of the LR axis. In the newborn and adult, expression of *lrd* mRNA was found in several types of ciliated epithelium, including the oviducts (Fig. 4d).

Expression of *lrd* mRNA in the embryo was observed primarily in non-ciliated cells, such as ES cells and blastocysts; it is unlikely that expression in the node is due to the presence of cilia, because cilia

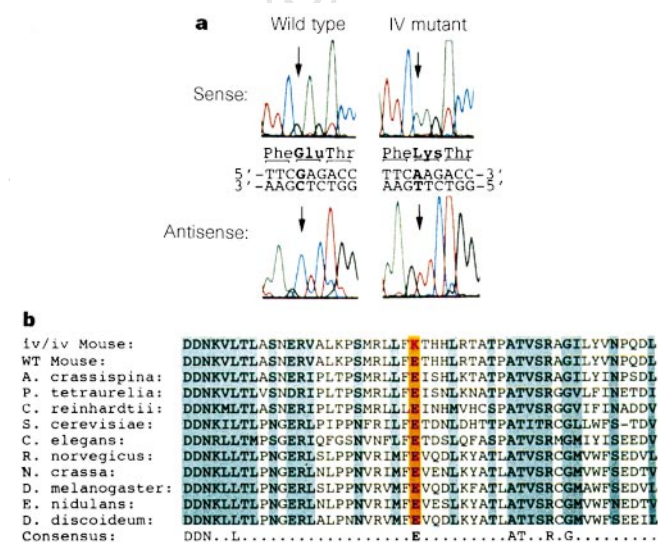


Figure 2 *iv* has a missense mutation in a highly conserved amino acid of *Lrd*. **a**, Raw sequence data. **b**, The E → K substitution is unique to *iv* and is in a very highly conserved amino-acid residue (orange). Amino acids that are identical or show only conservative changes across species are shaded. The consensus (bottom) shows residues that are identical among all known axonemal and cytoplasmic dynein sequences, including those not shown here. Amino-acid sequences in this region of *Lrd* were identical for all wild-type mouse strains analysed (12 in total). The rat orthologue also contains a glutamic acid at this position (data not shown).

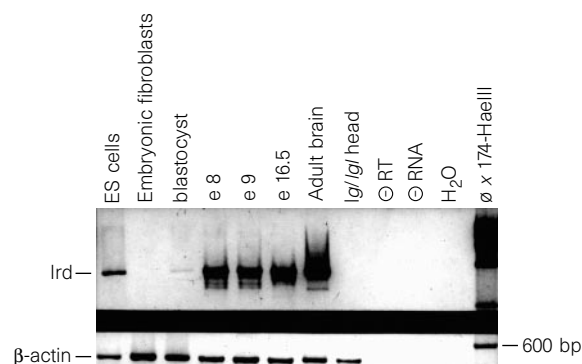
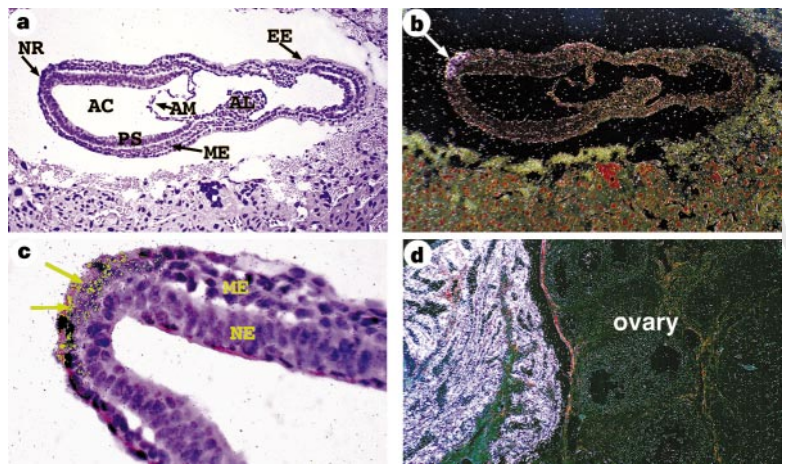


Figure 3 RT-PCR analysis of embryonic and adult expression. Primers used were specific for *lrd* (top) or control β -actin (bottom). Negative controls include: no reverse transcriptase (-RT), no tissue (medium control, -RNA), and no RNA (H₂O).

Figure 4 Expression of *lrd* localized by *in situ* hybridization. **a–d**, Hybridization to antisense *lrd* riboprobe; no signal was detected with a sense control probe. **a, c**, Bright-field image; **b, d**, dark-field images. **a, b**, Sagittal section of an E7.5 embryo demonstrating *lrd* expression in the node (arrow in **b**). **c**, Enlargement to show nodal localization of signal (arrows). **d**, *lrd* is expressed in the ciliated epithelium of the adult oviduct (left), but not in the ovary. NR, nodal region; AC, amniotic cavity; PS, primitive streak; AM, amnion; ME, mesoderm; AL, allantois; EE, extraembryonic ectoderm; NE, neural ectoderm. Computer-enhanced colour (yellow) was used to highlight the hybridization signal in **c**.



found in the node are immotile monocilia lacking dynein arms¹⁸. *lrd* mRNA expression was detected in many ciliated cells in the newborn and adult as well (Fig. 4d, and data not shown). As the sequence of *lrd* is more similar to axonemal than cytoplasmic dyneins, it is paradoxical that it is also found in cells lacking cilia; however, dyneins classified as axonemal have been found in non-ciliated cells¹⁹, suggesting that there is some blurring of the boundary between axonemal and cytoplasmic dyneins. We propose that *lrd* has dual functions: in the ciliated epithelium of adult structures it functions in ciliary beating, whereas embryonic expression indicates that mechanisms other than ciliary movement are involved in LR specification.

As a component of the dynein motor complex, *lrd* is an excellent candidate for involvement in LR determination, based on previous studies of dynein function and models of the development of handed asymmetry. As a microtubule-based motor, dynein might provide an initial LR bias to the embryo by asymmetric movement along a microtubule scaffolding which is oriented relative to the anterior–posterior and dorsal–ventral axes. Organized microtubule arrays are required for axis development in several model organisms^{20–22}. In *Xenopus*, for example, disruption of microtubule arrays in early embryos prevents cortical rotation and results in diminished dorsal–anterior development and randomized LR cardiac asymmetries²². Additionally, RNA encoding Vgl, a protein that can regulate *Xenopus* LR development²³, is asymmetrically localized to the vegetal cortex of the embryo through association with microtubules²⁴. These observations demonstrate the importance of microtubules in axis determination and the asymmetric localization of putative morphogens. Furthermore, the expression of *lrd* in the node is consistent with a role in LR specification, because several of the molecules implicated in LR development, such as *nodal*, *HNF-3 β* and *shh*, are also expressed there. A model of LR development²⁵ proposed the existence of handed molecules that are oriented with respect to the anterior–posterior and dorsal–ventral axes of the early embryo; these molecules would establish LR asymmetry by directional movements of morphogens. Based on our results showing that an axonemal dynein is mutated in *iv*, the handed molecules may be microtubules and dynein the molecular motor required to generate LR asymmetry. □

Methods

Mice. The *lgl* mutation has been described.³ SI/Col-*iv/iv* mice were obtained from Jackson Labs. For embryo isolation, noon on the day of finding a copulation plug (after overnight breeding) was considered as embryonic day 0.5. **YAC cloning.** F3, B4 and B6 YACs (Fig. 1a) were isolated as described²⁶. B14, B24 and B25 YACs were isolated from the Princeton YAC library screened with B4 YAC sequences (5′-GAAGAGCAGTCGGTGCTTTTACCAG-3′ and 5′-CAGACAGTTGTGAGCCACTGGATATA-3′), and C3 was isolated from the Baylor mouse YAC library using the same primers. YAC end-specific fragments

were isolated as described²⁶, and interspecific backcross chromosome mapping (Jackson Labs Backcross DNA Panel Map Service) was used to show that B14, B24 and B25 are chimaeric; the right end of C3 contains repetitive sequences and was therefore not mapped. Probes from the left end of F3, the right end of B4, and all non-repetitive probes isolated from C3 are deleted in *lgl/lgl* DNA.

***lrd* gene isolation.** A 270-bp dynein P-loop probe was cloned from mouse brain cDNA with primers (5′-AACCTTGTGATTCCAGGC-3′ and 5′-ATGGT-GATGAAGATAGCC-3′) based on conserved sequences flanking the first P-loop of the human and rat cytoplasmic dynein genes (GenBank accession numbers L23958 and L08505). This probe hybridized at reduced stringency (62°) to a 2.1-kb *Bgl*II fragment in C3, B14, B24 and B25 YAC DNAs. This cross-hybridizing fragment was cloned from a C3 YAC plasmid library and was used to screen adult mouse brain and testis cDNA libraries (Clontech) to isolate dyn-2 and dyn-22.2 cDNA clones.

The C3 YAC was used to clone additional cDNAs by direct cDNA selection as described²⁷, with the following modifications. cDNA prepared from wild-type adult brain mRNA using oligo(dT) and random primers and the Marathon cDNA-amplification kit (Clontech) was digested with *Alu*I, *Hae*III or *Rsa*I before linker addition and PCR amplification. cDNA was preblocked with 2 μ g Cot1 DNA (Gibco BRL) and 8 μ g *lgl/lgl* DNA to duplex repetitive and non-deleted sequences before hybridization. After two rounds of hybridization, selected cDNAs were subcloned into the pT7blue vector (Novagen).

RT-PCR using TaqStart Ab (Clontech) and Expand long-range PCR (Boehringer-Mannheim) was used to clone cDNAs connecting the library and direct selection cDNA sequences and to clone cDNAs from *iv/iv* mutant adult-brain RNA.

Chromosome mapping. Recombination mapping was done using the outcross/backcross panel previously described¹⁷. An 800-bp *Bst*XI restriction fragment probe from dyn-2 cDNA detected a polymorphism between SI/Col and C57BL/6J DNA digested with *Pst*I, allowing the C57BL/6J portion of the backcross panel to be used.

Mutation analysis. The region of *lrd* encompassing the mutation in *iv* was amplified using the following primers: dyn3.6 (5′-ATGATAA-CAAGGTGCTGACCC-3′), dyn3.7 (5′-CAAAATGCCAGCTCTGGAGACC-3′), dyn 3.6A (5′-GCCAGCAATGAACGAGTGGCCCTCAAACCT-3′), and dyn 3.7A (5′-AGCTCTGGAACCGTGGCTGGTGTGGCTGT-3′). Primers dyn3.6 and dyn3.7 were used for RT-PCR to clone fragments from wild-type (Outbred ICR mice, Harlan) and *iv/iv* adult brain RNA. Three different RT-PCR clones per genotype, representing two different RNA samples, were analysed. For confirmation, these primers were used to isolate PCR products from genomic DNA of wild-type mouse strains DBA, 129/J, C57BL/6J, C3H, SJL/J and *Mus spretus*, as well as *iv/iv* (SI/Col), for direct sequencing with primers dyn3.7 and dyn3.8 (genomic-specific: 5′-TGAGTCTACCTGTGGCT-GTCC-3′). Additionally, primers dyn3.6A and dyn3.7A were used to clone PCR fragments from SI/Col-*iv/iv* and the wild-type strains (C57BL/6J, C3H, BALB/c, A/J, AKR/J, SPRET/Ei (*Mus spretus*), PANCEVO/Ei (*Mus hortulanus*), TIRANO/Ei (*Mus domesticus*) and rat (*Rattus norvegicus*). At least two independent clones per DNA were sequenced. DNAs were obtained from the Jackson Labs.

GenBank accession numbers for sequences shown in Fig. 2b: *Anthocardis crassispina* (sea urchin) ciliary dynein- β heavy chain (D01021), *Paramecium tetraurelia* outer arm dynein- β heavy chain (U19464), *Chlamydomonas reinhardtii* dynein- β heavy chain (U02963), *Saccharomyces cerevisiae* cytoplasmic dynein heavy chain (Z21877), *Caenorhabditis elegans* dynein (Z75536), *Rattus norvegicus* cytoplasmic dynein heavy chain (L08505), *Neurospora crassa* cytoplasmic dynein heavy chain (P45443), *Drosophila melanogaster* cytoplasmic dynein heavy chain (P37276), *Emmericella nidulans* cytoplasmic dynein (P45444), and *Dictyostelium discoideum* cytoplasmic dynein heavy chain (A44357).

RT-PCR. cDNA was prepared using SuperScript (Gibco BRL). PCR was done using AmpliTaq Gold (Perkin-Elmer) and primers dyn3.3 (5'-AGAGCTT-CATCCTCAAAGTTGTC-3') and dyn3.5 (5'-CAACTGGCCACATACG-GATTC-3') (94 °C for 30 s, 65 °C for 30 s, 72 °C for 30 s; 42 cycles).

In situ hybridization. Serial-section *in situ* hybridization was performed as described²⁸ with a cDNA probe from a non-conserved region of *lrd* (the most 3' 694 nucleotides of the sequence indicated in Fig. 1b); this probe does not cross-hybridize with other dynein genes by Southern blot analysis.

Received 15 July; accepted 16 September 1997.

- Hummel, K. P. & Chapman, D. B. Visceral inversion and associated anomalies in the mouse. *J. Hered.* **50**, 9–13 (1959).
- Layton, W. M. Random determination of a developmental process. *J. Hered.* **67**, 336–338 (1976).
- McNeish, J. D. *et al.* Phenotypic characterization of the transgenic mouse insertional mutation *Legless*. *J. Exp. Zool.* **253**, 151–162 (1990).
- Singh, G. *et al.* *legless* insertional mutation: morphological, molecular, and genetic characterization. *Genes Dev.* **5**, 2245–2255 (1991).
- Lowe, L. A. *et al.* Conserved left-right asymmetry of nodal expression and alterations in murine *situs inversus*. *Nature* **381**, 158–161 (1996).
- Meno, C. *et al.* Left-right asymmetric expression of the TGF β -family member *lefty* in mouse embryos. *Nature* **381**, 151–155 (1996).
- Lohr, J. L., Danos, M. C. & Yost, H. J. Left-right asymmetry of a *nodal*-related gene is regulated by dorsoanterior midline structures during *Xenopus* development. *Development* **124**, 1465–1472 (1997).
- Afzelius, B. A. A human syndrome caused by immotile cilia. *Science* **193**, 317–319 (1976).
- Afzelius, B. A. *Situs inversus* and ciliary abnormalities: What is the connection? *Int. J. Dev. Biol.* **39**, 839–844 (1995).
- Holzbaur, E. L. F. & Vallee, R. B. Dyneins: molecular structure and cellular function. *Annu. Rev. Cell Biol.* **10**, 339–372 (1994).
- Asai, D. J. Multi-dynein hypothesis. *Cell Motil. Cytoskel.* **32**, 129–132 (1995).
- Tanaka, Y., Zhang, Z. & Hirokawa, N. Identification and molecular evolution of new dynein-like protein sequences in rat brain. *J. Cell Sci.* **108**, 1883–1893 (1995).
- Vaughan, K. T. *et al.* Multiple mouse chromosomal loci for dynein-based motility. *Genomics* **36**, 29–38 (1996).
- Andrews, K. L., Nettesheim, P., Asai, D. J. & Ostrowski, L. E. Identification of seven rat axonemal dynein heavy chain genes: expression during ciliated cell differentiation. *Mol. Biol. Cell* **7**, 71–79 (1996).
- McGrath, J., Horwich, A. L. & Brueckner, M. Duplication/deficiency mapping of *situs inversus viscerum* (*iv*), a gene that determines left-right asymmetry in the mouse. *Genomics* **14**, 643–648 (1992).
- Asai, D. J. *et al.* The dynein genes of *Paramecium tetraurelia*: sequences adjacent to the catalytic P-loop identify cytoplasmic and axonemal heavy chain isoforms. *J. Cell Sci.* **107**, 839–847 (1994).
- Brueckner, M., D'Eustachio, P. & Horwich, A. L. Linkage mapping of a mouse gene, *iv*, that controls left-right asymmetry of the heart and viscera. *Proc. Natl Acad. Sci. USA* **86**, 5035–5038 (1989).
- Bellomo, D., Lander, A., Harragan, I. & Brown, N. A. Cell proliferation in mammalian gastrulation: the ventral node and notochord are relatively quiescent. *Dev. Dyn.* **205**, 471–485 (1996).
- Vaisberg, E. A., Grissom, P. M. & McIntosh, J. R. Mammalian cells express three distinct dynein heavy chains that are localized to different cytoplasmic organelles. *J. Cell Biol.* **133**, 831–841 (1996).
- Theurkauf, W. E. Microtubules and cytoplasm organization during *Drosophila* oogenesis. *Dev. Biol.* **165**, 352–360 (1994).
- Jesuthasan, S. & Strahle, U. Dynamic microtubules and specification of the zebrafish embryonic axis. *Curr. Biol.* **7**, 31–42 (1996).
- Danos, M. C. & Yost, H. J. Linkage of cardiac left-right asymmetry and dorsal-anterior development in *Xenopus*. *Development* **121**, 1467–1474 (1995).
- Hyatt, B. A., Lohr, J. L. & Yost, H. J. Initiation of vertebrate left-right axis formation by maternal *Vg1*. *Nature* **384**, 62–65 (1996).
- Elisha, Z., Havin, L., Ringel, I. & Yisraeli, J. K. *Vg1* RNA binding protein mediates the association of *Vg1* RNA with microtubules in *Xenopus* oocytes. *EMBO J.* **14**, 5109–5114 (1995).
- Brown, N. A., McCarthy, A. & Wolpert, L. Development of handed body asymmetry in mammals. *Ciba Found. Symp.* **162**, 182–201 (1991).
- Supp, D. M., Witte, D. P., Branford, W. W., Smith, E. P. & Potter, S. S. Sp4, a member of the Sp1-family of zinc finger transcription factors, is required for normal murine growth, viability, and male fertility. *Dev. Biol.* **176**, 284–299 (1996).
- Segre, J. A., Nemhauser, J. L., Taylor, B. A., Nadeau, J. H. & Lander, E. S. Positional cloning of the *nude* locus: genetic, physical, and transcription maps of the region and mutations in the mouse and rat. *Genomics* **28**, 549–559 (1995).
- Kern, M. J., Witte, D. P., Valerius, M. T., Aronow, B. J. & Potter, S. S. A novel murine homeobox gene isolated by a tissue specific PCR cloning strategy. *Nucleic Acids Res.* **20**, 5189–5195 (1992).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from Mary Sheehan at the London editorial office of Nature.

Acknowledgements. We thank R. B. Vallee, M. A. Ghee and K. T. Vaughan for their comments and for providing a mouse cytoplasmic dynein cDNA probe; J. McGrath for assistance in embryo isolation; M. Lee for help with chromosome mapping; J. M. Corrales for help with cDNA cloning; K. Saalfeld, P. Groen and L. Artmeyer for *in situ* hybridization; A. Emley for photographic assistance; S. Bell for help with the expression studies; A. Horwich for comments and discussion; and the staff at the Keck Biotechnology Center at Yale University and at the University of Cincinnati DNA Core Facility for DNA sequencing. This work was supported by grants from the NIH (S.S.P. and D.M.S.), the American Heart Association, Ohio Division (D.P.W.), and The March of Dimes (M.B.).

Correspondence and requests for materials should be addressed to M.B. (e-mail: brueckner@biomed.med.yale.edu).

Wnt signalling required for expansion of neural crest and CNS progenitors

Makoto Ikeya*, Scott M. K. Lee†, Jane E. Johnson‡, Andrew P. McMahon† & Shinji Takada*

* Centre for Molecular and Developmental Biology, Faculty of Science, Kyoto University, Kitashirakawa, Sakyo-ku, Kyoto 606-01, Japan

† Department of Molecular and Cellular Biology, The Biolabs,

Harvard University, 16 Divinity Avenue, Cambridge, Massachusetts 02138, USA

‡ Cell Biology and Neuroscience, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, Texas 75235, USA

Interactions between cells help to elaborate pattern within the vertebrate central nervous system (CNS)¹. The genes *Wnt-1* and *Wnt-3a*, which encode members of the Wnt family of cysteine-rich secreted signals, are coexpressed at the dorsal midline of the developing neural tube, coincident with dorsal patterning^{2,3}. Each signal is essential for embryonic development, *Wnt-1* for mid-brain patterning^{4,5} and *Wnt-3a* for formation of the paraxial mesoderm⁶, but the absence of a dorsal neural-tube phenotype in each mutant suggests that Wnt signalling may be redundant. Here we demonstrate that in the absence of both *Wnt-1* and *Wnt-3a* there is a marked deficiency in neural crest derivatives, which originate from the dorsal neural tube⁷, and a pronounced reduction in dorsolateral neural precursors within the neural tube itself. These phenotypes do not seem to result from a disruption in the mechanisms responsible for establishing normal dorsoventral polarity. Rather, our results are consistent with a model in which local Wnt signalling regulates the expansion of dorsal neural precursors. Given the widespread expression of different *Wnt* genes in discrete areas of the mammalian neural tube³, this may represent a general model for the action of Wnt signalling in the developing CNS.

The Wnt family of cysteine-rich signalling factors is composed of at least 16 members in the mouse. Four of these, *Wnt-1*, -3, -3a and -4, are expressed in largely overlapping regions within the dorsal CNS, predominantly at the dorsal midline from the forebrain to spinal cord³. Neural expression of the genes encoding *Wnt-1* and *Wnt-3a* occurs before that of other family members, coincident with neural-crest differentiation in the dorsal neural tube. Expression remains in the roofplate throughout the period of CNS neurogenesis^{2,3}. Null mutants in three of these *Wnt* genes have been described: *Wnt-1* regulates midbrain development^{4,5}, *Wnt-3a* regulates paraxial mesoderm formation⁶, and *Wnt-4* regulates kidney tubulogenesis⁸. None of these signals seems to be required for dorsal CNS development. However, cell transformation studies indicate that *Wnt-1* and *Wnt-3a* have similar activities⁹. Thus redundancy between Wnt signals may obscure a possible role for Wnt signalling in formation of the neural crest and in dorsoventral patterning of the vertebrate CNS.

To address this possibility we generated compound mutants lacking both *Wnt-1* and *Wnt-3a* signalling. Compound homozygotes were recovered at the expected mendelian frequency (51 compound homozygotes in 673 embryos, close to the expected frequency of 1/16) between 9.0 and 10.5 days post coitum (d.p.c.). However, owing to the termination of caudal axial development accompanying the loss of *Wnt-3a* activity, relatively few of these embryos survived to 18.5 d.p.c. (3 compound homozygotes in 151 embryos).

To assess the development of the dorsal neural tube in compound mutants we first examined structures derived from the neural crest. Neural crest cells are amongst the first differentiated cell types to be

formed by dorsal neural precursors. Neurofilament staining indicated that formation of both neural crest-derived cranial and spinal ganglia were unaltered in single mutants that were either wild type or heterozygous for mutations in the other *Wnt* member (Fig. 1a–f). However, in double mutants, neurons derived from the proximal ganglion of cranial nerve IX (glossopharyngeal), which is formed by crest cells originating from rhombomere 6 (r6) within the hindbrain^{10,11}, were absent (Fig. 1g). In contrast, the distal ganglion, which is placodal in origin, was present. There was also a marked reduction in the proximal axons of cranial nerves V (trigeminal, r2 derived) and X (vagus, r7 derived). Similarly, in the trunk there was a dramatic reduction in neurofilament staining in the cervical dorsal root ganglia (Fig. 1h), and cell counts indicated a 60% decrease in the cellular content of the dorsal root ganglia. Whole-mount *in situ* hybridization with probes recognizing *Islet-1* (Fig. 2a–d) and *cadherin-6* (ref. 12) (Fig. 2e–h), which are markers of neuronal and glial neural crest derivatives, respectively, within the cranial ganglia and dorsal root ganglia confirmed the reduction or absence of crest cells. In contrast, sympathetic ganglia, which express *c-ret*¹³, were unaffected (Fig. 2i, j).

The reduction of neurogenic and gliogenic crest derivatives in the caudal head and rostral trunk regions suggests that fewer neural crest cells may emerge in embryos lacking both *Wnt-1* and *Wnt-3a*

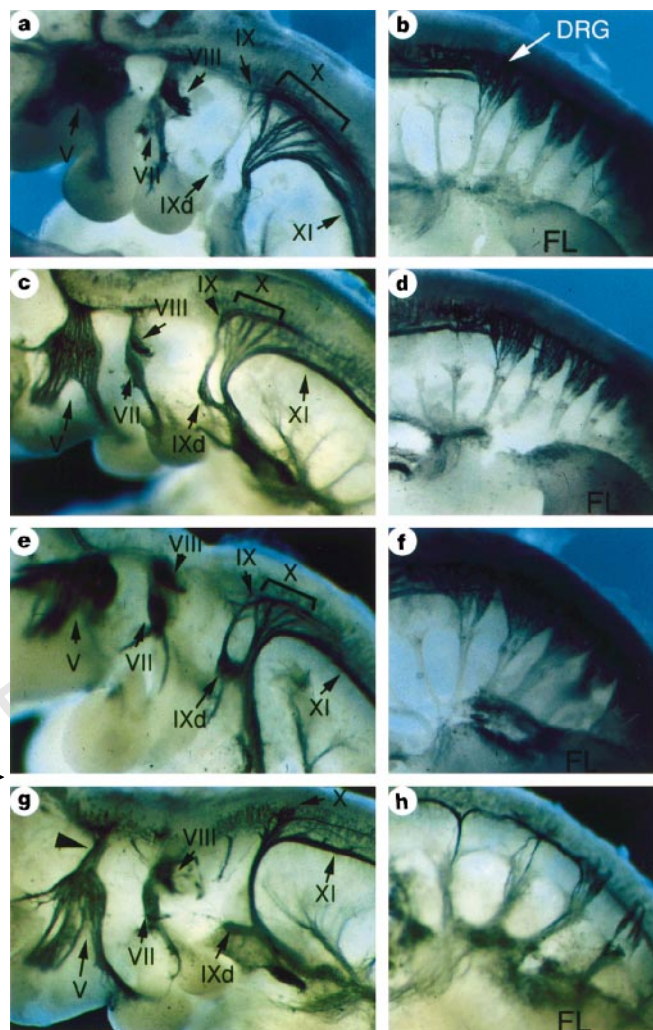


Figure 1 Analysis of neuronal patterning in *Wnt-1/Wnt-3a* compound mutants. Neurons in 10.5 d.p.c. embryos were visualized by whole-mount immunostaining using an anti-neurofilament antibody. Cranial region (**a, c, e, g**) and trunk just rostral to the forelimb levels (**b, d, f, h**) of wild-type (**a, b**), *Wnt-1* homozygous (**c, d**), *Wnt-3a* homozygous (**e, f**) and *Wnt-1* and *Wnt-3a* compound homozygous (**g, h**) embryos. Cranial nerves are indicated as follows: trigeminal ganglia, V; facial ganglia, VII; acoustic ganglia, VIII; glossopharyngeal nerve, IX; distal ganglia of glossopharyngeal nerve, IXd; vagus nerve, X; and accessory nerve, XI. An arrowhead in **g** indicates reduction in the proximal axon derived from trigeminal ganglia. Dorsal root ganglia (DRG) and forelimbs (FL) are also indicated.

Figure 2 Analysis of the expression of neuronal and glial markers in *Wnt-1/Wnt-3a* compound mutants. **a–d**, *Islet-1* expression in the cranial region (**a, b**) and rostral trunk (**c, d**) of a *Wnt-1* homozygous (**a**), wild-type (**c**) and compound homozygous (**b, d**) embryos at 10.5 d.p.c. **e–h**, *Cadherin-6* expression in the cranial region (**e, f**) and rostral trunk (**g, h**) of wild-type (**e, g**) and compound homozygous (**f, h**) embryos at 10.5 d.p.c. **i, j**, *c-ret* whole-mount *in situ* hybridization to 10.5 d.p.c. wild-type (**i**) and compound mutant (**j**) embryos. In **a, b, e** and **f**, cranial nerves are indicated as in Fig. 1. Arrowheads in **c, d, g** and **h** indicate DRG. Forelimbs (FL) and sympathetic ganglia (SG) are indicated.



Figure 3 Analysis of the expression of neural crest cell markers in *Wnt-1/Wnt-3a* compound mutants. **a, b**, CRABP-1 immunostaining in 9.0 d.p.c. wild-type (**a**) and compound mutant (**b**) embryos. Compound mutants display a loss of anterior brain regions in addition to the loss of midbrain and rhombomere 1 of the hindbrain, which is typical of *Wnt-1* single mutants. This aspect of the phenotype will be reported elsewhere. **c-h**, *AP-2* expression in the cranial (**c, d**) and rostral trunk region (**e-h**) of wild-type (**c, e, g**) and compound mutant (**d, f, h**) embryos. Transverse sections (**c, d, g** and **h**) and dorsal views (**e, f**) of whole-mount stained embryos are shown. **i-l**, lateral (**i, j**) and dorsal (**k, l**) views of *TRP-2* expression in 11.5 d.p.c. wild-type (**i, k**) and compound mutant (**j, l**) embryos. **m**, A higher magnification of the marked area in **l**, showing that a few *TRP-2* positive cells remained at the dorsal midline. Arrows in **c-h** indicate migrating neural crest cells. Asterisks in **i-l** indicate hindbrain region. Second (r2), fourth (r4), sixth rhombomeres (r6) and forelimbs (FL) are indicated.

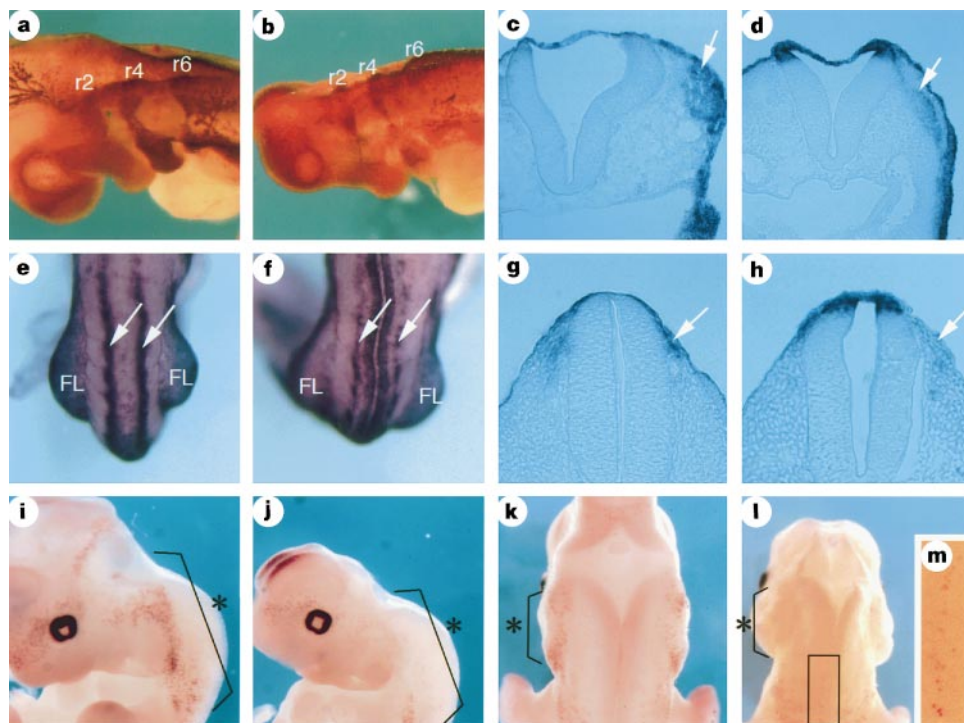
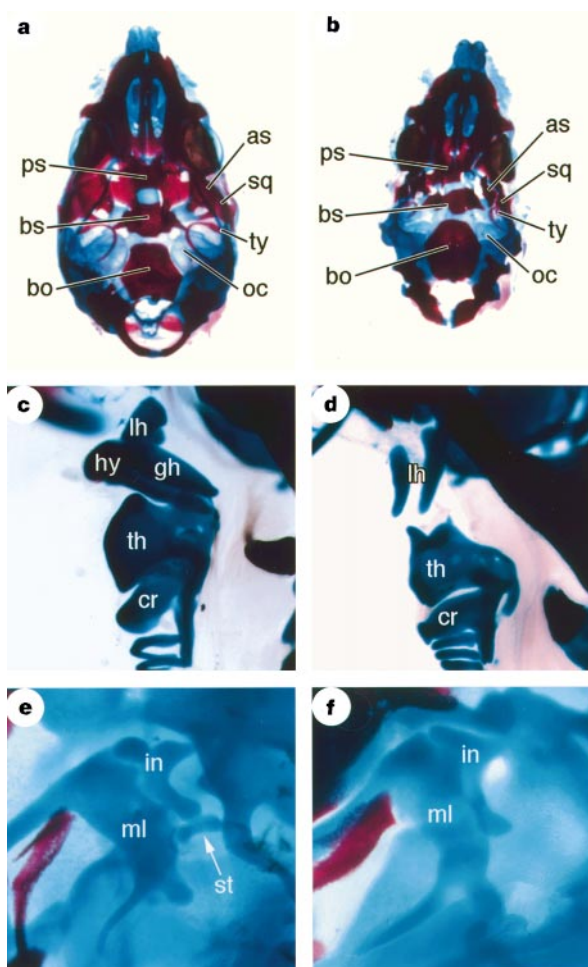


Figure 4 Cranial skeleton of *Wnt-1/Wnt-3a* compound mutant embryos at 18.5 d.p.c. Cranial skeletal elements were unaltered in either single mutant except for a part of the parietal bone, a non-crest derivative, which was missing in *Wnt-1* homozygotes. **a, b**, Ventral views of the cranial skeleton of wild-type (**a**) and compound mutant (**b**) embryos. **c-f**, Lateral views of skeleton preparation of the laryngeal cartilages (**c, d**) and the middle ear ossicles cartilages (**e, f**) in wild-type (**c, e**) and compound mutant (**d, f**) embryos. Presphenoid, ps; basisphenoid, bs; basioccipital, bo; alisphenoid, as; squamosal, sq; tympanic, ty; otic capsule, oc; hyoid bone, hy; greater horn of hyoid bone, gh; lesser horn of hyoid bone, lh; thyroid cartilage, th; cricoid cartilage, cr; malleus, ml; incus, in; stapes, st.



signalling. The issue of neural crest formation was addressed by examining CRABP-1 immunoreactivity and *AP-2* transcription. CRABP-1 is normally present in the dorsal CNS at 9.0 d.p.c. as well as in migrating neural crest cells arising from r2, r4 and r6 (ref. 14). *AP-2* is first expressed at 8.5 d.p.c. in the dorsal neural plate, coincident with neural crest formation. By 9.5 d.p.c. cranial expression is absent in the neural tube but persists in migrating and maturing neural crest derivatives at cranial and spinal cord levels¹⁵. Loss-of-function studies have demonstrated that *AP-2* is essential for the development of neural crest-derived structures^{16,17}. A clear decrease was observed in migrating cells positive for CRABP-1 within the hindbrain (Fig. 3a, b), although CRABP-1 staining within the CNS was relatively normal. Similarly, examination of *AP-2* expression revealed a reduction in both cranial (Fig. 3c, d) and trunk (Fig. 3e, f) neural crest. Surprisingly, in contrast to their wild-type littermates, double mutants also retained *AP-2* expression within the dorsal CNS at 9.5 d.p.c. (Fig. 3c, d, g, h). Thus, in the absence of *Wnt-1* and *Wnt-3a*, there is both a reduction in neural crest cell formation and persistent expression of *AP-2* at the dorsal midline.

To determine whether Wnt signalling was required throughout

the period of cranial crest formation, we examined expression of *TRP-2*, a marker of presumptive melanocytes¹⁸, which are dominant in late-formed cranial crest derivatives¹⁹. At 11.5 d.p.c., *TRP-2* expression was virtually absent within presumptive melanocytes migrating within the hindbrain region of double mutants, although a few *TRP-2* cells remained at the dorsal midline (Fig. 3i–m). Because of the prolonged expression of *AP-2* within the dorsal CNS, *TRP-2* expressing cells may still be differentiating at later stages. Alternatively, these cells may be retained at the midline because Wnt signalling promotes neural crest migration. *In vitro* assays that examined neural crest induction by surface ectoderm found a correlation between the appearance of *Wnt-1* expression and migration of neural crest cells²⁰. Expression of *CRABP-1* (Fig. 3a, b), *TRP-2* (Fig. 3i, j) and *AP-2* (data not shown) remained constant in the forebrain, suggesting that there was some regional specificity in the requirement for these Wnt signals.

Much of the head skeleton is generated by cranial neural crest. Distinct skeletal elements are derived from neural crest cells that emerge from different regions of the brain^{7,21,22}. To determine whether the reduction in neural crest formation in double mutants led to alterations in the skeleton, embryos at 18.5 d.p.c. were stained with alcian blue and alizarin red to examine the development of cartilage and bone (Fig. 4a–f). The stapes and the main body of the hyoid bone including the greater horn, which originate from crest cells derived from r3–5 and r6–7, respectively, were absent, and the

thyroid cartilage showed a consistent dysmorphology (Fig. 4d, f). Thus the reduction in hindbrain crest formation is also reflected in the absence of specific skeletal derivatives. In contrast, despite the early loss of forebrain, midbrain and rostral hindbrain in double mutants (S.M.K.L., M.I., S.T. & A.P.M., manuscript in preparation), the development of skeletal crest derivatives for these regions, as well as non-crest-derived bones, was largely normal, although there was some reduction in the development of the squamosal, alisphenoid, basisphenoid, presphenoid and otic capsule (Fig. 4e). In summary, in the absence of Wnt-1 and Wnt-3a signalling, several neural crest cell fates form, but there is a reduction in neural crest derivatives in both the hindbrain and spinal cord.

The development of neural crest cells and other aspects of dorsal polarity within the developing CNS are thought to be regulated by signals from bone morphogenetic proteins (BMPs) produced initially by the dorsal ectoderm and subsequently by the dorsal CNS²³. Expression of *Bmp-7* was induced, as expected, in the roofplate of double mutants (data not shown). Thus it was unlikely that defective neural crest development resulted from a secondary loss of BMP signalling within the dorsal neural tube. To determine whether Wnt signalling may directly regulate dorsoventral polarity within the CNS we examined the distribution of several regionally expressed markers. Levels of the markers in the spinal cord appeared normal (data not shown), but the hindbrain displayed a striking phenotype. Expression of *Wnt-3a* (Fig. 5a–d), *Wnt-1* and *Lmx-1b*

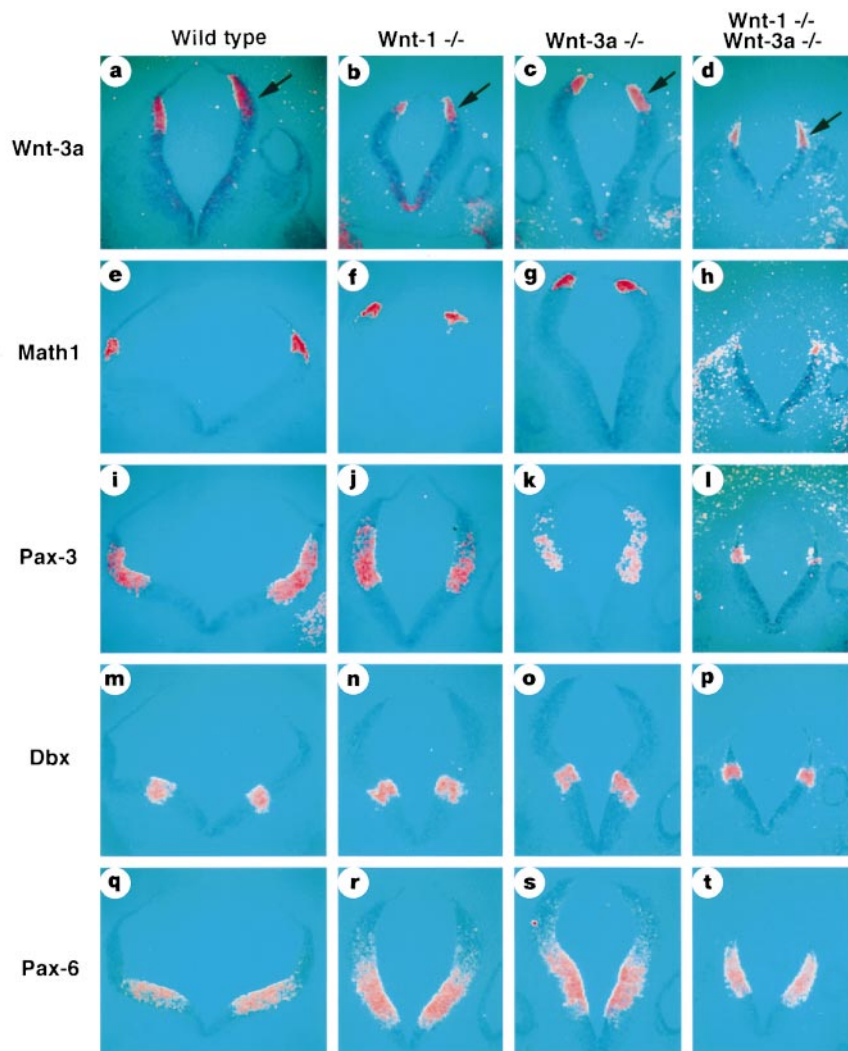


Figure 5 Analysis of dorsal-ventral polarity in the hindbrain of *Wnt-1/Wnt-3a* compound mutants. *In situ* hybridization at the otic level with the probes and genotypes as indicated. Arrows indicate the position of the roofplate.

(data not shown) was slightly reduced but localized, as normal, to the roofplate. In contrast, expression of *Math1*, which is not activated until 9.5 d.p.c. in cells that overlap and extend lateral to the *Wnt-1* expression domain (Fig. 5e–h), and *Pax-3*, which occupies most of the dorsal half of the CNS (Fig. 5i–l), were reduced dramatically in the double-mutant hindbrain. Expression of *Dbx* at the dorsal–ventral interface (Fig. 5m–p) and *Pax-6* in the ventrolateral hindbrain (Fig. 5q–t) were normal.

Our data suggest that, in the hindbrain, Wnt signalling is not involved in primary patterning processes, as cells expressing markers of distinct dorsoventral fates occupy appropriate positions along the dorsoventral axis. Rather, there is a pronounced reduction in dorsolateral progenitors, suggesting that Wnt signalling regulates the expansion of these cells, as either a survival factor or a mitogen. The latter explanation is supported by transgenic analysis, which demonstrated that *Wnt-1* can act as a potent mitogen when expressed ectopically within the dorsal CNS²⁴. Further, a proliferative model would explain the dorsoventral progression in the neural crest phenotype we observed. In normal development there is a ventral to dorsal progression in the formation of different neural crest derivatives^{25,26}. In the double mutants, the most severely affected crest derivatives were more proximal (dorsally located) structures; for instance, the stapes was absent from the second branchial arch but the lesser horn of the hyoid was unaltered; and in the trunk, dorsal root ganglia were markedly reduced, although the sympathetic ganglia appeared normal. If the signals governing commitment to neural crest cell fates were unperturbed in the double mutant, but renewal of a multipotential dorsal neural progenitor pool required Wnt signals, the expected result would be a loss of later-forming neural crest derivatives in *Wnt-1* and *Wnt-3a* mutants, as precursors within the neural tube became limiting. Further, if a certain number of cell divisions are required before *AP-2* expression is downregulated in the dorsal neural tube, a similar mechanism might also explain the persistence of *AP-2* expression in the hindbrain of double mutants.

We investigated cell proliferation and cell death by using bromodeoxyuridine incorporation and TUNEL staining, respectively, within hindbrain sections at 9.5 d.p.c. (20–25 somites). At this stage there was already a dramatic reduction of dorsal neural precursors, but no discernible change was detected in either proliferation or cell death within remaining dorsal regions of *Wnt-1* and *Wnt-3a* mutants (data not shown). As these two *Wnt* genes are first coexpressed at the otic level when the first neural crest cells appear, at around 8.5 d.p.c. (8–10 somites)^{3,25,26}, it is likely that the main reduction in dorsolateral neural precursors occurs between 8.5 and 9.5 d.p.c.

Taken together with earlier transgenic analysis, our results favour a model in which localized Wnt signalling regulates dorsoventral patterning in the mammalian CNS through the control of cell proliferation. Many members of the Wnt family are expressed regionally within the developing mammalian CNS. Whether the local regulation of cell proliferation represents a general mechanism applicable to other Wnt signalling pathways remains to be determined, however. □

Methods

Generation and collection of embryos. Generation of mice heterozygous for null alleles of *Wnt-1* and *Wnt-3a* was described previously^{4,6}. Compound heterozygotes (on a predominantly 129/Sv background) were intercrossed to recover compound mutants. Genotypes were initially confirmed by genomic Southern hybridization and PCR.

Whole-mount immunostaining. Whole-mount neurofilament staining²⁷ was visualized using a monoclonal antibody (2H3) from the Developmental Hybridoma Bank at the NICHD. CRABP-1 immunostaining was performed on 9.0 d.p.c. embryos using a polyclonal antibody provided by U. Eriksson through N. Ohsumi. Lmx-1b immunostaining was performed at 10.5 d.p.c. using a rabbit polyclonal antiserum (A.P.M., unpublished data).

In situ hybridization. Whole-mount digoxigenin³ and ³⁵S section *in situ* hybridization² was performed using the probes indicated.

Skeleton preparations. Preparation of skeletons was performed according to published procedure²⁸. Briefly, 18.5 d.p.c. embryos were eviscerated, skinned, fixed in 80% alcohol, and stained with alcian blue and alizarin red.

Cell proliferation and cell death. Embryos were collected at 9.5 d.p.c. (20–25 somite stage equivalent) after intraperitoneal injection of pregnant females with 50 µg per g body weight of 5-bromo-2'-deoxyuridine (BrdU), and the mice were killed 1 h later. Embryos were fixed overnight in 4% paraformaldehyde in phosphate-buffered saline at 4 °C. After dehydration, wax embedding and sectioning at a thickness of 6 µm, serial sections were dewaxed and either stained with haematoxylin and eosin, or assayed for BrdU incorporation²⁹ or for apoptotic cell death using the TUNEL procedure³⁰.

Received 13 January; accepted 1 September 1997.

1. Tanabe, Y. & Jessell, T. M. Diversity and pattern in the developing spinal cord. *Science* **274**, 1115–1123 (1996).
2. Wilkinson, D. G., Bailes, J. A. & McMahon, A. P. Expression of the proto-oncogene *int-1* is restricted to specific neural cells in the developing mouse embryo. *Cell* **50**, 79–88 (1987).
3. Parr, B. A., Shea, M. J., Vassileva, G. & McMahon, A. P. Mouse *Wnt* genes exhibit discrete domains of expression in the early embryonic CNS and limb buds. *Development* **119**, 247–261 (1993).
4. McMahon, A. P. & Bradley, A. The *Wnt-1* (*int-1*) proto-oncogene is required for development of a large region of the mouse brain. *Cell* **62**, 1073–1085 (1990).
5. Thomas, K. R. & Capecchi, M. R. Targeted disruption of the murine *int-1* proto-oncogene resulting in severe abnormalities in midbrain and cerebellar development. *Nature* **346**, 847–850 (1990).
6. Takada, S. *et al.* *Wnt-3a* regulates somite and tailbud formation in the mouse embryo. *Genes Dev.* **8**, 174–189 (1994).
7. Le Douarin, N. M. *The Neural Crest* (Cambridge Univ. Press, 1982).
8. Stark, K., Vainio, S., Vassileva, G. & McMahon, A. P. Epithelial transformation of metanephric mesenchyme in the developing kidney regulated by *Wnt-4*. *Nature* **372**, 679–683 (1994).
9. Wong, G. T., Gavin, B. J. & McMahon, A. P. Differential transformation of mammary epithelial cells by *Wnt* genes. *Mol. Cell. Biol.* **14**, 6278–6286 (1994).
10. D'Amico-Martel, A. & Noden, D. M. Contributions of placodal and neural crest cells to avian cranial peripheral ganglia. *Am. J. Anat.* **166**, 445–468 (1983).
11. Lumsden, A., Sprawson, N. & Graham, A. Segmental origin and migration of neural crest cells in the hindbrain region of the chick embryo. *Development* **113**, 1281–1291 (1991).
12. Inoue, T., Chisaka, O., Matsunami, H. & Takeichi, M. Cadherin-6 expression transiently delineates specific rhombomeres, other neural tube subdivisions and neural crest subpopulations in mouse embryos. *Dev. Biol.* **183**, 183–194 (1997).
13. Pachnis, V., Mankoo, B. & Costantini, F. Expression of the *c-ret* proto-oncogene during mouse embryogenesis. *Development* **119**, 1005–1017 (1993).
14. Maden, M. *et al.* Domains of cellular retinoic acid-binding protein I (CRABP I) expression in the hindbrain and neural crest of the mouse embryo. *Mech. Dev.* **37**, 13–23 (1992).
15. Chazaud, C. *et al.* *AP-2.2*, a novel gene related to *AP-2*, is expressed in the forebrain, limbs and face during mouse embryogenesis. *Mech. Dev.* **54**, 83–94 (1995).
16. Zhang, J. *et al.* Neural tube, skeletal and body wall defects in mice lacking transcription factor *AP-2*. *Nature* **381**, 238–241 (1996).
17. Schorle, H. *et al.* Transcription factor *AP-2* essential for cranial closure and craniofacial development. *Nature* **381**, 235–238 (1996).
18. Wehrle-Haller, B. & Weston, J. A. Soluble and cell-bound forms of steel factor activity play distinct roles in melanocyte precursor dispersal and survival on the lateral neural crest migration pathway. *Development* **121**, 731–742 (1995).
19. Serbedzija, G. N., Bronner-Fraser, M. & Fraser, S. E. A vital dye analysis of the timing and pathways of avian neural crest cell migration. *Development* **106**, 809–816 (1989).
20. Dickinson, M. E. *et al.* Dorsalization of the neural tube by non-neural ectoderm. *Development* **121**, 2099–2106 (1995).
21. Köntges, G. & Lumsden, A. Rhombencephalic neural crest segmentation is preserved throughout craniofacial ontogeny. *Development* **122**, 3229–3242 (1996).
22. Couly, G., Grapin-Botton, A., Coltey, P. M. & LeDouarin, N. M. The regeneration of the cephalic neural crest, a problem revisited: the regenerating cells originate from the contralateral or from the anterior and posterior neural fold. *Development* **122**, 3393–3407 (1996).
23. Leim, K. F., Tremml, G., Roelink, H. & Jessell, T. M. Dorsal differentiation of neural plate cells induced by BMP-mediated signals from epidermal ectoderm. *Cell* **82**, 969–979 (1995).
24. Dickinson, M. E., Krumlauf, R. & McMahon, A. P. Evidence for a mitogenic effect of *Wnt-1* in the developing mammalian central nervous system. *Development* **120**, 1453–1471 (1994).
25. Serbedzija, G. N., Fraser, S. E. & Bronner-Fraser, M. Pathways of trunk neural crest cell migration in the mouse embryo as revealed by vital dye labelling. *Development* **108**, 605–612 (1990).
26. Serbedzija, G. N., Bronner-Fraser, M. & Fraser, S. E. Vital dye analysis of cranial neural crest cell migration in the mouse embryo. *Development* **116**, 297–307 (1992).
27. Mark, M. *et al.* Two rhombomeres are altered in *Hoxa-1* mutant mice. *Development* **119**, 319–338 (1993).
28. Parr, B. A. & McMahon, A. P. Dorsalizing signal *Wnt-7a* required for normal polarity of D-V and A-P axes of mouse limb. *Nature* **374**, 350–353 (1995).
29. Miller, M. W. & Nowakowski, R. S. Use of bromodeoxyuridine immunohistochemistry to examine the proliferation, migration and time of origin of cells in the central nervous system. *Brain Res.* **457**, 44–52 (1988).
30. Gavrieli, Y., Sherman, Y. & Ben-Sasson, S. A. Identification of programmed cell death in situ via specific labeling of nuclear DNA fragmentation. *J. Cell Biol.* **119**, 493–501 (1992).

Acknowledgements. We thank M. Takeichi for discussion and the *cadherin-6* complementary DNA probe; U. Eriksson and N. Ohsumi for the CRABP-1 polyclonal antibody; T. Jessell for the *Islet-1* cDNA probe; A. Dudley for the *Bmp-7* cDNA probe; T. Williams for the *AP-2* cDNA probe; M. Takahashi for the *c-ret* cDNA probe; P. Gruss for *Pax-3* and *Pax-6* cDNA probes; and I. Jackson for the *TRP-2* cDNA probe. This work was supported by a grant-in-aid for Creative Fundamental Research from the Ministry of Education, Science, and Culture of Japan (S.T.), a grant from the Kent Waldrep National Paralysis Foundation (J.E.J.) and a grant from the NIH (A.P.M.).

Correspondence and requests for materials should be addressed to either S.T. (e-mail: stakada@take.biophys.kyoto-u.ac.jp) or A.P.M. (amcmahon@biosun.harvard.edu).

Ataxin-1 with an expanded glutamine tract alters nuclear matrix-associated structures

Pamela J. Skinner*, Beena T. Koshy†, Christopher J. Cummings‡, Ivan A. Klement*, Kara Helin*, Antonio Servadio§☆, Huda Y. Zoghbi†§|| & Harry T. Orr*¶#

* Departments of Laboratory Medicine and Pathology and † Biochemistry, and # Institute of Human Genetics, University of Minnesota, Minneapolis, Minnesota 55455, USA

† Department of Pediatrics, ‡ Program in Cell and Molecular Biology, and

§ Department of Molecular and Human Genetics, Baylor College of Medicine, and || Howard Hughes Medical Institute, Houston, Texas 77030, USA

Spinocerebellar ataxia type 1 (SCA1) is one of several neurodegenerative disorders caused by an expansion of a polyglutamine tract^{1,2}. It is characterized by ataxia, progressive motor deterioration, and loss of cerebellar Purkinje cells¹. To understand the pathogenesis of SCA1, we examined the subcellular localization of wild-type human ataxin-1 (the protein encoded by the *SCA1* gene) and mutant ataxin-1 in the Purkinje cells of transgenic mice³. We found that ataxin-1 localizes to the nuclei of cerebellar Purkinje cells. Normal ataxin-1 localizes to several nuclear structures ~0.5 µm across, whereas the expanded ataxin-1 localizes to a single ~2-µm structure, before the onset of ataxia. Mutant ataxin-1 localizes to a single nuclear structure in affected neurons of SCA1 patients. Similarly, COS-1 cells transfected with wild-type or mutant ataxin-1 show a similar pattern of nuclear localization; with expanded ataxin-1 occurring in larger structures that are fewer in number than those of normal ataxin-1. Colocalization studies show that mutant ataxin-1 causes a specific redistribution of the nuclear matrix-associated domain containing promyelocytic leukaemia protein⁴⁻⁷. Nuclear matrix preparations demonstrate that ataxin-1 associates with the nuclear matrix in Purkinje and COS cells. We therefore propose that a critical aspect of SCA1 pathogenesis involves the disruption of a nuclear matrix-associated domain.

To investigate the pathogenesis of the polyglutamine disorder SCA1, we examined the subcellular localization of wild-type human ataxin-1 and mutant ataxin-1 in the Purkinje cells of transgenic mice³. Because the ataxin-1 antibodies we used recognize both murine and human ataxin-1 (ref. 8), we used *SCA1* transgenic mice³ crossed onto a murine ataxin-1 null background⁹. Ataxin-1 localized to the nuclei, but not the nucleoli, of Purkinje cells in mice carrying either a wild-type (30 repeat) or mutant (82 repeat) *SCA1* allele at all ages examined (Fig. 1a, b). In Purkinje cells of transgenic mice expressing a wild-type *SCA1* allele, ataxin-1 localized throughout the nucleus and to multiple nuclear structures that were ~0.5 µm in size (Fig. 1c). Mutant ataxin-1 also localized throughout the nucleus but, in contrast to wild-type ataxin-1, it localized to a single large ~2-µm nuclear structure (Fig. 1d). The fraction of Purkinje cells containing the large nuclear inclusion of mutant ataxin-1 increased with age from 25% at six weeks of age to 90% at 12 weeks. Additionally, cells expressing mutant ataxin-1 often demonstrated invaginations in the nuclear membrane. The age of onset of ataxia, determined by cage behaviour for these mice, is 18 weeks (P.J.S. *et al.*, unpublished data); thus the changes in nuclear structure develop before the onset of ataxia.

We used immunohistochemistry to examine the subcellular

distribution of ataxin-1 in brain tissue from several human SCA1 patients. We found ubiquitin-positive nuclear inclusions of ataxin-1 in the nucleus pontis centralis, a region affected by the disease¹⁰. As an example, the results in Fig. 1e, f are from one patient with 29 and 82 CAG repeats. Examination of two other brain regions affected by SCA1, the olives and Purkinje cells, revealed substantial cell loss, making the identification of ataxin-1-positive neurons almost impossible. However, in the rare instances when a neuron could be identified in the inferior olive, the presence of ataxin-1 nuclear inclusions was demonstrated. Nuclear inclusions of ataxin-1 were not detected in brain regions unaffected by the disease. The results from SCA1 transgenic mice and patients demonstrate that ataxin-1 localizes to distinct nuclear domains in neurons. Furthermore, they demonstrate that an expanded form of ataxin-1 has a nuclear distribution that differs from that of wild-type ataxin-1.

In view of the nuclear localization of ataxin-1 and the nuclear changes induced by mutant ataxin-1 in Purkinje cells of transgenic mice and affected neurons of SCA1 patients, we examined the nuclear localization of ataxin-1 in a transfected cell line. We used monkey kidney COS-1 cells, because they do not express endogenous ataxin-1 yet readily express ataxin-1 upon transfection with *SCA1* constructs (Fig. 2a). When either a wild-type or expanded form of ataxin-1 was expressed in transfected COS cells, there was no evidence that ataxin-1 induced cell death. In these cells, trans-

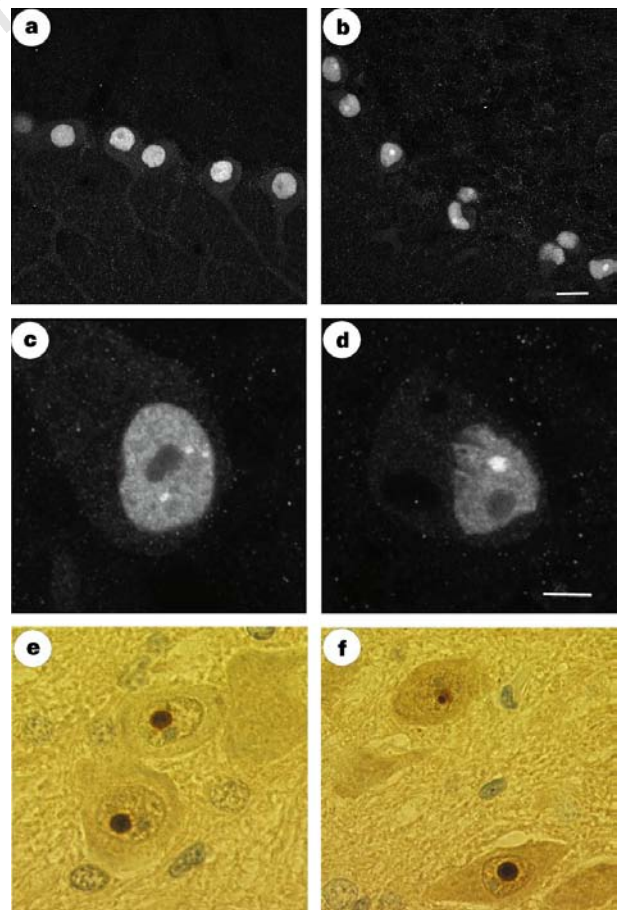


Figure 1 Immunofluorescence of ataxin-1 within the nuclei of transgenic mouse Purkinje cells and SCA1 patient neurons. **a, c**, Localization of wild-type ataxin-1 with 30 repeats; **b, d**, localization of mutant ataxin-1 with 82 repeats in the Purkinje cells of transgenic mice. In **a** and **b**, a 10-µm series of confocal images are projected (scale bar, 15 µm). In **c** and **d** a single confocal Z scan is shown (scale bar, 5 µm). **e, f**, Ataxin-1 (**e**) and ubiquitin (**f**) immunostaining of neurons is shown in the nucleus pontis centralis from a SCA1 patient with an expanded allele of 82 repeats.

* Present address: Department of Biotechnologies, Telethon Institute of Genetics and Medicine, Milan, Italy.

fecting wild-type as well as mutant ataxin-1 localized to the nucleus and were excluded from the nucleoli (Fig. 2b–g). Most (75%) of the cells transfected with normal ataxin-1 showed a diffuse staining accompanied by many small structures of less than 1 μm in diameter (Fig. 2d). In contrast, 60% of COS-1 cells transfected with mutant ataxin-1 had ataxin-1-containing nuclear structures that were fewer in number and larger in size, sometimes exceeding 5 μm in diameter (Fig. 2g). Western blot analysis revealed that levels of mutant and wild-type ataxin-1 were similar (data not shown), suggesting that the difference in nuclear distribution between wild-type and mutant ataxin-1 was not due to a difference in expression levels. We conclude that the altered nuclear distribution of mutant ataxin-1 is due to the expanded polyglutamine tract.

There is increasing evidence that the nucleus is highly organized into functional domains¹¹, so we used molecular probes for selected nuclear components to characterize further the nuclear distribution of ataxin-1. Promyelocytic leukaemia protein (PML) is a nuclear matrix-associated protein and a component of PML oncogenic domains (PODs), which are also known as nuclear bodies^{4–7}. We used the antibodies to PML, SC35 (a component of spliceosomes)¹², p80-coilin (a component of coiled bodies)¹³ and BCL-6 (a transcription factor)¹⁴, in colocalization studies of SCA1-transfected COS-1 cells. Both wild-type and mutant ataxin-1 failed to colocalize with SC35, p80-coilin or BCL-6, and they had no effect on the

nuclear distribution of these proteins (Fig. 3d–l). However, in cells double-labelled with anti-ataxin-1 (ref. 8) and anti-PML antibodies⁴, the large nuclear structures characteristic of mutant ataxin-1 often sequestered PML and clearly altered its normal nuclear distribution (Fig. 3a–c). Although small structures of ataxin-1 occasionally colocalized with PML in COS cells transfected with wild-type ataxin-1, these are likely to happen by chance given the high number of structures formed by each protein per cell and the fact that most of the PML and ataxin-1 structures are distinct. Most importantly, expression of wild-type ataxin-1 in COS cells did not alter the normal distribution of PML. Overexpression of BCL-6 in COS cells also had no effect on PML nuclear distribution¹⁴, suggesting that the disruption of PML by mutant ataxin-1 is specific. Furthermore, the fact that the nuclear distribution of three other nuclear components was unaffected by the expression of mutant ataxin-1 suggests that the formation of the large structures containing ataxin-1 and PML is unlikely to be the result of a general loss of nuclear structural health and integrity.

Because PML is known to associate with the nuclear matrix, we investigated whether ataxin-1 is also a nuclear matrix-associated protein. We isolated the nuclear matrix from COS-1 cells transfected with ataxin-1 by means of cell lysis with a non-ionic detergent, nucleic acid digestion, and high salt extraction¹⁵. When the resultant matrix was stained with propidium iodide and anti-PML and anti-ataxin-1 antibodies, we detected PML and both wild-type (Fig. 4a) and mutant ataxin-1 (Fig. 4b). Ataxin-1 was also detected in western blot analysis of nuclear matrix preparations of transfected COS-7 cells (Fig. 4c). To demonstrate that ataxin-1 is a component of the nuclear matrix in cerebellar tissue, nuclear matrix preparations were performed on brain tissue from a wild-type SCA1 transgenic mouse. Western blot analysis of this material revealed the presence of ataxin-1 (Fig. 4d).

Our analyses of the subcellular localization of ataxin-1 in the Purkinje cells of SCA1 transgenic mice, neurons of SCA1 patients, and in transfected COS-1 cells have revealed that ataxin-1 localizes to nuclear compartments that associated with the nuclear matrix. In these systems the pattern of nuclear distribution differed for mutant and wild-type ataxin-1, as mutant ataxin-1 alone was associated with specific alterations of nuclear structure. These data suggest that the subcellular target of SCA1 pathogenesis is within the nucleus and involves the disruption of a nuclear function. Histological examination of neurons from individuals affected with two other polyglutamine disorders, Huntington disease and SCA3, revealed the presence of single large nuclear inclusions^{16–19} similar to those reported here. Furthermore, transgenic mice expressing a fragment of the Huntington's disease gene with an expanded number of glutamines have nuclear inclusions containing the protein huntingtin²⁰. Thus polyglutamine disorders involve pathogenic changes in the nucleus. Our observations that ataxin-1 associates with the nuclear matrix and that mutant ataxin-1 causes the specific redistribution of PML, another matrix-associated protein, may implicate a specific nuclear mechanism. The neuronal specificity of SCA1 pathogenesis is likely to be determined by the interaction of ataxin-1 with other nuclear proteins unique to those neurons affected by SCA1, such as the Purkinje-cell protein called leucine-rich acidic nuclear protein (LANP)²¹. □

Methods

Immunohistochemical analysis of cerebella. Anaesthetized mice were perfused with PBS-buffered formalin, fixed overnight, placed in PBS and stored at 4 °C. Vibratome sections ~50 μm thick were cut in the sagittal plane. Ataxin-1 epitopes were revealed by microwaving the sections three times for 20 s in 0.01 M urea^{8,22}. The sections were blocked for 1 h using 2% normal goat sera in PBS with 0.3% Triton X-100 and then incubated for 48 h in blocking solution with a 1:400 to 1:1,000 dilution of anti-ataxin 11750VII antibodies or a 1:4,000 to 1:10,000 dilution of anti-ataxin-1 11750VII (ref. 9) at 4 °C. Sections were washed four times for 20 min in PBS, incubated with blocking solution

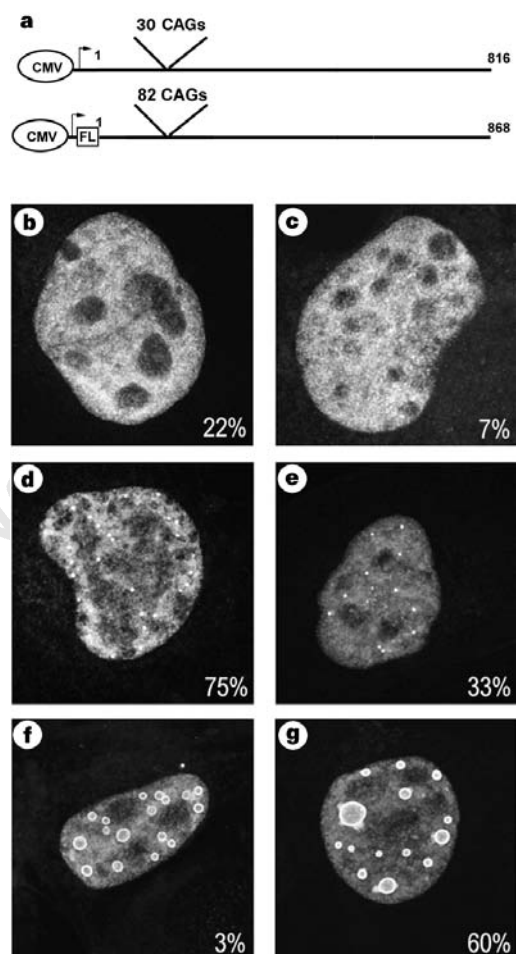


Figure 2 Immunofluorescence of ataxin-1 in transfected COS-1 cells. **a**, The constructs. CMV, the CMV promoter; FL, the Flag epitope; numbers, amino-acid residues. **b, d, f**, The subcellular localization of wild-type ataxin-1[30]; **c, e, g**, the localization of mutant ataxin-1[82]. The percent of transfected cells with each depicted pattern it indicated. Images were generated by projecting a series of confocal Z scans that spanned the nucleus.

Figure 3 Confocal analysis of ataxin-1[82] and other nuclear proteins in COS-1 cells. **a**, Endogenous PML; **d**, endogenous SC35; **g**, endogenous p80-coilin; **j**, transfected BCL-6; **b**, **e**, **h**, **k**, mutant ataxin-1[82]; **c**, **f**, **i**, **l**, Merged images of PML and ataxin-1[82] (**c**), SC35 and ataxin-1[82] (**f**), p80-coilin and ataxin-1[82] (**i**), and BCL-6 and ataxin-1 (**l**). Scale bar, 5 μ m.

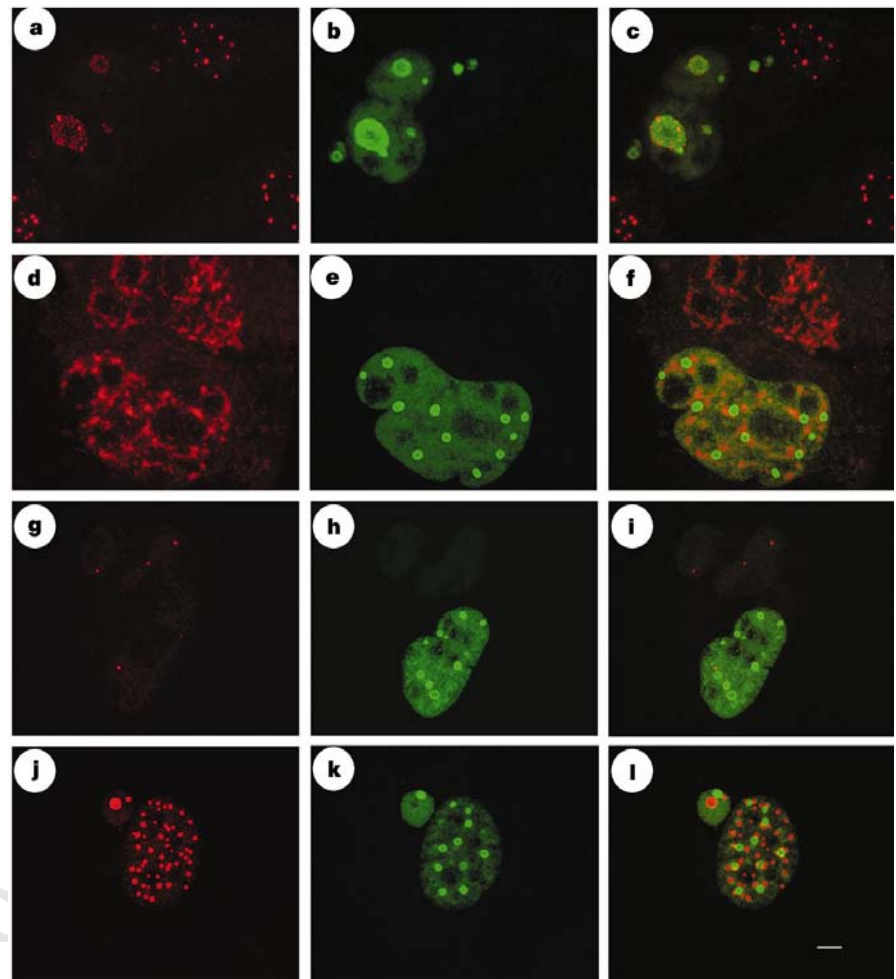
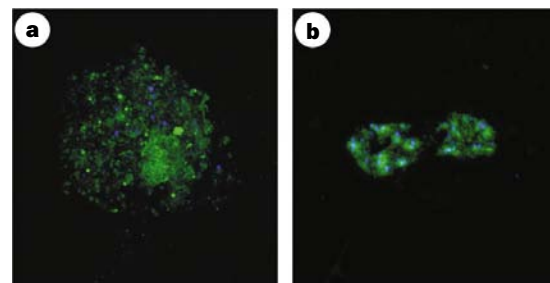
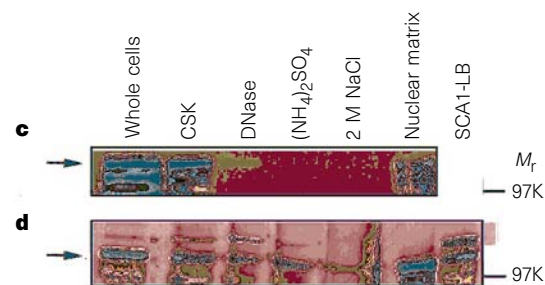


Figure 4 Association of ataxin-1 with the nuclear matrix. **a**, **b**, Ataxin-1[30] (**a**) and ataxin-1[82] (**b**) (green) in nuclear matrix from transfected COS-1 cells that were counterstained with propidium iodide (red) as a negative control, and anti-PML (blue), as a positive control. **c**, **d**, Western blots of subcellular fractions from ataxin-1[82] transfected COS-7 cells (**c**) and cerebella of ataxin-1[30] transgenic mice (**d**), probed with anti-ataxin-1 sera. The arrow indicates the ataxin-1 band. The subcellular fractions analysed in **c** and **d** include whole-cell protein extract; CSK extracts from whole cells (**c**) and isolated nuclei (**d**); DNase-soluble protein, DNase; $(\text{NH}_4)_2\text{SO}_4$ -soluble protein, $(\text{NH}_4)_2\text{SO}_4$; NaCl-soluble protein, 2 M NaCl; and protein remaining with the nuclear matrix. A lymphoblast (LB) lysate from an SCA1 patient, having alleles of 26 and 82 CAGs, was loaded into the last lane of **d**.



containing goat anti-rabbit polysera conjugated to Cy3 (Jackson ImmunoResearch) for 48 h, and washed four times for 20 min in PBS. Sections were mounted on slides with glycerol-gelatin containing 4 mg ml⁻¹ *n*-propylgalate and examined with a Bio-Rad MRC-1000 confocal microscope equipped with a krypton-argon laser. Human brain tissue from a SCA1 patient (29 and 82 CAG repeats) was obtained 2 h after death and immediately fixed. Paraffin sections (5 μ m) were used for immunostaining⁹. Antibodies used were an anti-ataxin-1 antiserum (11NQ) and a monoclonal anti-ubiquitin antibody (Novacastra). The 11NQ antiserum was raised to an N-terminal peptide (amino acids 164–197) of ataxin-1 conjugated through its N terminus to keyhole limpet haemocyanin. The 11NQ antiserum was raised in New Zealand rabbits by Cocalico Biologicals and characterized (A.S. & H.Y.Z., unpublished data).

Plasmids. The SCA1 cDNAs^{3,23} were subcloned into the *Eco*RI site of the eukaryotic expression vector pcDNA1Amp (Invitrogen). A Flag epitope was inserted downstream of the ATG codon in the SCA1[82] construct. Plasmid pVB140(EFBT), from T. Miki (Tokyo Medical and Dental University), contains a Myc-tagged BCL-6 cDNA under the control of the elongation factor promoter.



Immunohistochemical analysis of transfected COS-1 cells. The DEAE-dextran method of transfection²⁴ was used to transfect 1×10^6 COS-1 cells in 60-mm tissue-culture plates. Cells were transferred to sterile coverslips in 6-well tissue-culture plates 24 h post-transfection. Cells were fixed for 10 min in PBS plus 3.7% formaldehyde 48 h post-transfection, and permeabilized for 2 min with cold acetone. Coverslips were coated with 40 μ l of primary antibody

diluted in PBS. Either 1:100 dilutions of anti-flag M5 monoclonal antibodies (Kodak), anti-ataxin-1 11750V (1:100) or anti-ataxin-1 11750VII (1:200) polysera⁸ were used as primary antibodies. Anti-SC35 hybridoma supernatant¹² was used undiluted. Anti-PML (1:200) polysera⁵ and anti-PML monoclonal antibody 5E10 (1:10 dilution)⁴ were used. The anti-Myc Tag monoclonal antibody 9E10 (from V. Bardwell, University of Minnesota) was used to detect BCL-6. Anti-p80-coilin antibody^{25,26} was used to visualize coiled bodies. Goat anti-mouse conjugated to FITC (Caltag), goat anti-rabbit conjugated to lissamine rhodamine (Accurate Chemical and Scientific), goat-anti rabbit and goat anti-mouse conjugated to Cy2, Cy3 or Cy5 (Jackson ImmunoResearch) were used as secondary antibodies. Propidium iodide diluted in PBS (1 µg ml⁻¹) was used to stain nucleic acid. A subset of cells was also stained with bisbenzimidazole (0.001%) 10 min RT (Hoechst 33258 Sigma) as a transfection control. Images were collected and processed as described above.

Preparation of nuclear matrix. COS-1 cells seeded on coverslips were used for nuclear matrix preparations 48 h after transfection, as described previously¹⁵. Isolated matrices on coverslips were stained by immunofluorescence as described above. Nuclear matrix preparations from transgenic mice cerebella and COS-7 cells were carried out as described^{27–29}. Immunoblotting was performed as described⁸. Nuclei were extracted with SK buffer (10 mM PIPES, pH 6.8, 300 mM sucrose, 100 mM NaCl, 3 mM MgCl₂, 1 mM EGTA, 0.5% Triton X-100, 1.2 mM PMSF, 2 mM vanadyladenosine and protease inhibitors) for 3 min at 4°C. Pelleted nuclei were suspended in digestion buffer (SCK plus 50 mM NaCl) with 400 U ml⁻¹ of RNase-free DNase I and digested for 50 min at 32°C. Digested DNA was extracted with 0.25 M (NH₄)₂SO₄ and washed with 2 M NaCl in digestion buffer. For COS-7 cells, 2 × 10⁶ cells were extracted analogous to nuclear matrix preparation from transgenic mice, except that extractions were carried out on whole cells instead of isolated nuclei.

Received 9 July; accepted 8 September 1997.

- Zoghbi, H. Y. & Orr, H. T. Spinocerebellar ataxia type 1. *Semin. Cell Biol.* **6**, 29–35 (1995).
- Ross, C. A. When more is less: Pathogenesis of glutamine repeat neurodegenerative diseases. *Neuron* **15**, 493–496 (1995).
- Burridge, E. N. et al. SCA1 transgenic mice: a model for neurodegeneration caused by an expanded CAG trinucleotide repeat. *Cell* **82**, 937–948 (1995).
- Stuurman, N. et al. A monoclonal antibody recognizing nuclear matrix-associated nuclear bodies. *J. Cell Sci.* **101**, 773–784 (1992).
- Weis, K. et al. Retinoic acid regulates aberrant nuclear localization of PML-RAR alpha in acute promyelocytic leukemia cells. *Cell* **76**, 345–356 (1994).
- Koken, M. H. et al. The t(15;17) translocation alters a nuclear body in a retinoic acid-reversible fashion. *EMBO J.* **13**, 1073–1083 (1994).
- Dyck, J. A. et al. A novel macromolecular structure is a target of the promyelocyte-retinoic acid receptor oncoprotein. *Cell* **76**, 333–343 (1994).
- Servadio, A. et al. Expression analysis of the ataxin-1 protein in tissues from normal and spinocerebellar ataxia type 1 individuals. *Nature Genet.* **10**, 94–98 (1995).
- Matilla, A. et al. Ataxin-1, the SCA1 gene product, is required for learning tasks mediated by both the hippocampus and cerebellum. *Proc. Natl Acad. Sci. USA* (submitted).
- Robitaille, Y., Schut, L. & Kish, S. J. Structural and immunocytochemical features of olivopontocerebellar atrophy caused by the spinocerebellar ataxia type 1 (SCA-1) mutation define a unique phenotype. *Acta Neuropathol.* **90**, 572–581 (1995).
- Spector, D. L. Macromolecular domains within the cell nucleus. *Annu. Rev. Cell Biol.* **9**, 265–315 (1993).
- Fu, X. D. & Maniatis, T. Factor required for mammalian spliceosome assembly is localized to discrete regions in the nucleus. *Nature* **343**, 437–441 (1990).
- Raska, I. et al. Immunological and ultrastructural studies of the nuclear coiled body with autoimmune antibodies. *Exp. Cell Res.* **195**, 27–37 (1991).
- Dhordain, P. et al. The BTB/POZ domain targets the LAZ3/BCL6 oncoprotein to nuclear dots and mediates homomerization in vivo. *Oncogene* **11**, 2689–2697 (1995).
- Bisotto, S., Lauriault, P., Duval, M. & Vincent, M. Colocalization of a high molecular mass phosphoprotein of the nuclear matrix (p255) with spliceosomes. *J. Cell Sci.* **108**, 1873–1882 (1995).
- Roizin, S. S. & Liu, J. C. Neuronal nuclear-cytoplasmic changes in Huntington's Chorea: electron microscope investigations. *Adv. Neurol.* **23**, 95–122 (1979).
- Tellez-Nagel, I., Johnson, A. B. & Terry, R. D. Studies on brain biopsies of patients with Huntington's chorea. *J. Neuropathol. Exp. Neurol.* **33**, 308–332 (1974).
- Roos, R. A. C. & Bots, G. T. A. M. Nuclear membrane indentations in Huntington's Chorea. *J. Neurol. Sci.* **61**, 37–47 (1983).
- Paulson, H. L. et al. Intranuclear inclusions of expanded polyglutamine protein in spinocerebellar ataxia type 3. *Neuron* **19**, 333–344 (1997).
- Davies, S. W. et al. Formation of neuronal intranuclear inclusions (NII) underlies the neurological dysfunction in mice transgenic for the HD mutation. *Cell* **90**, 537–548 (1997).
- Matilla, A. et al. The cerebellar leucine-rich acidic nuclear protein interacts with ataxin-1. *Nature* **389**, 974–978 (1997).
- Cattorelli, G. et al. Antigen unmasking on formalin-fixed, paraffin-embedded tissue sections. *J. Pathol.* **171**, 83–98 (1993).
- Banfi, S. et al. Identification and characterization of the gene causing type 1 spinocerebellar ataxia. *Nature Genet.* **7**, 513–520 (1994).
- Ausubel, F. M. et al. in *Current Protocols in Molecular Biology* 9.2.1–9.2.3 (John Wiley, New York, 1996).
- Andrade, L. E., Tan, E. M. & Chan, E. K. Immunocytochemical analysis of the coiled body in the cell cycle and during cell proliferation. *Proc. Natl Acad. Sci. USA* **90**, 1947–1951 (1993).

- Chan, E. K., Takano, S., Andrade, L. E., Hamel, J. C. & Matera, A. G. Structure, expression and chromosomal localization of human p80-coilin gene. *Nucleic Acids Res.* **22**, 4462–4469 (1994).
- Tawfik, S. & Ahmed, K. Association of casein kinase 2 with nuclear matrix. Possible role in nuclear matrix protein phosphorylation. *J. Biol. Chem.* **269**, 7489–7493 (1994).
- Fey, E. G. & Penman, S. Nuclear matrix proteins reflect cell type origin in cultured human cells. *Proc. Natl Acad. Sci. USA* **85**, 121–125 (1988).
- Berezney, R. & Coffey, D. S. Nuclear matrix isolation and characterization of a framework structure from rat liver nuclei. *J. Cell Biol.* **73**, 616–632 (1977).

Acknowledgements. We thank V. Bardwell for assistance with the BCL-6 transfections; E. N. Burridge for assistance with SCA1 transgenic mice; D. Saxon for assistance with tissue sectioning and staining; G. Sedgewick for assistance with confocal microscopy and image processing; D. Armstrong for assistance in analysis of SCA1 human tissue; T. Maniatis for the anti-SC35 antibody; A. Dejean for the anti-PML polysera; R. Van Driel for the anti-PML monoclonal antibody 5E10; and E. K. Chan for the anti-p80-coilin polysera. This work was supported by grants from the National Institute of Neurological Disorders and Stroke of the NIH to H.T.O. and H.Y.Z. H.Y.Z. is a Howard Hughes Medical Institute investigator.

Correspondence and requests for materials should be addressed to H.T.O. (e-mail: harry@lenti.med.umn.edu).

The cerebellar leucine-rich acidic nuclear protein interacts with ataxin-1

Antoni Matilla*, Beena T. Koshy*, Christopher J. Cummings*†, Toshiaki Isobe‡, Harry T. Orr§ & Huda Y. Zoghbi*†||

Departments of * Pediatrics, † Molecular and Human Genetics, and Program in Cell and Molecular Biology, Baylor College of Medicine, and || Howard Hughes Medical Institute, Houston, Texas 77030, USA

‡ Department of Chemistry, Faculty of Science, Tokyo Metropolitan University, Tokyo, Japan

§ Departments of Laboratory Medicine and Pathology, and Biochemistry, and Institute of Human Genetics, University of Minnesota, Minneapolis, Minnesota 55455, USA

Spinocerebellar ataxia type 1 (SCA1) is an autosomal dominant neurodegenerative disorder characterized by ataxia, progressive motor deterioration, and loss of cerebellar Purkinje cells¹. SCA1 belongs to a growing group of neurodegenerative disorders caused by expansion of CAG repeats, which encode glutamine². Although the proteins containing these repeats are widely expressed, the neurodegeneration in SCA1 and other polyglutamine diseases selectively involves a few neuronal subtypes. The mechanism(s) underlying this neuronal specificity is unknown. Here we show that the cerebellar leucine-rich acidic nuclear protein (LANP)³ interacts with ataxin-1, the SCA1 gene product. LANP is expressed predominantly in Purkinje cells, the primary site of pathology in SCA1. The interaction between LANP and ataxin-1 is significantly stronger when the number of glutamines is increased. Immunofluorescence studies demonstrate that both LANP and ataxin-1 colocalize in nuclear matrix-associated sub-nuclear structures. The features of the interaction between ataxin-1 and LANP, their spatial and temporal patterns of expression, and the colocalization studies indicate that cerebellar LANP is involved in the pathogenesis of SCA1.

The protein ataxin-1, the SCA1 gene product, is found predominantly in the nucleus in neurons and the cytoplasm in peripheral tissues⁴. In SCA1, there is selective degeneration of cerebellar Purkinje cells and brainstem neurons¹ despite the wide expression pattern of ataxin-1. Cell-specific proteins may mediate the pathogenesis of SCA1 and the other polyglutamine neurodegenerative diseases.

To identify proteins that interact with ataxin-1, we performed a yeast two-hybrid genetic screen⁵ using a mutant human ataxin-1 containing 82 glutamines. Yeast expressing mutant ataxin-1 was transformed with a mouse brain cDNA library, and screening of approximately 1 × 10⁷ tryptophan and leucine auxotrophic transformants resulted in the identification of four positive clones. Sequence analysis of these clones identified partial cDNAs encoding 147 amino acids homologous to the rat leucine-rich acidic nuclear

protein (rLANP; expressed in cerebellum and brainstem)³, ankyrin G (expressed at the nodes of Ranvier and axonal initial segments in brain, spinal cord and peripheral nerve)⁶, myo-inositol 1-phosphate synthase (expressed in brain vasculature and peripheral tissues)⁷, and the neurotransmitter transporter rB21a (expressed in leptomeninges)⁸. Because LANP was the only protein known to be expressed in tissues affected in SCA1, it was used for additional studies.

The specificity of the interaction between LANP and ataxin-1 was demonstrated by mating yeast containing LANP with a battery of yeast clones of the opposite mating type expressing 10 different proteins. Only yeast containing both LANP and ataxin-1 with either 30 glutamines or 82 glutamines activated both the *lacZ* and *HIS3* reporter genes.

A full-length mouse cDNA clone (*Lanp*) was isolated from a murine cerebellar cDNA library. Alignment of the deduced mouse LANP amino-acid sequence with those from rat and bovine LANP sequences revealed 98.4% and 92.7% amino-acid identity, respectively. To test the reproducibility of the interaction between LANP and ataxin-1, the full-length mouse *Lanp* cDNA was subcloned into the pAS2-CYH2 vector and the 82-glutamine ataxin-1[82] into the pACT2 vector. When these two clones were co-transformed into yeast, histidine prototrophy and strong β -galactosidase activity were detected.

To define the regions of ataxin-1 that interact with LANP, bait constructs containing different regions of the SCA1 cDNA were generated⁹ and transformed into yeast expressing *Lanp* cDNA. Activation of the *lacZ* reporter gene was noted when ataxin-1 subclones spanning the glutamine repeats, or the last 268 amino acids, were used (Fig. 1). However, the strongest interaction between the two proteins was obtained when full-length ataxin-1 containing 82 glutamines was used. In contrast, no interaction was detected when the first 195 amino acids or amino acids 277–579 of ataxin-1 were used as bait. These results demonstrate that LANP interacts with ataxin-1, and that the interaction is modulated by the number of glutamines. The fact that LANP also interacts with ataxin-1 outside the polyglutamine domain suggests that the interaction between LANP and ataxin-1 is relevant for the cellular function of both proteins. The portion of LANP interacting with ataxin-1 is the first 147 amino acids, containing five leucine-rich repeats (LRRs), an amphipathic motif known to mediate protein–protein interactions¹⁰.

To determine whether the strength of the interaction between

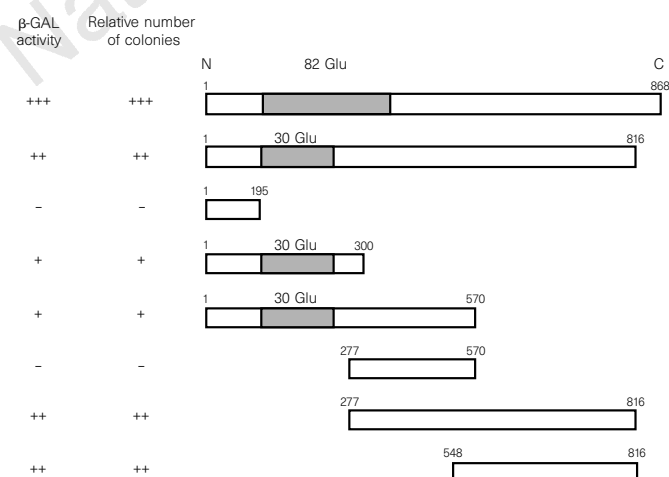


Figure 1 Schematic representation of constructs used in the yeast two-hybrid screen to assay interactions between ataxin-1 and LANP. Numbers indicate the positions of amino acids in ataxin-1. The polyglutamine tract and the C-terminal region (amino acids 570–816) of ataxin-1 interact with LANP. More colonies and a higher level of β -galactosidase activity, as detected by intensity of blue colour, are observed with the full-length mutant ataxin-1[82].

ataxin-1 and LANP was influenced by the number of glutamines, we assayed β -gal activity using two different approaches. First, we assayed the sensitivity of the interaction using filter lifts of yeast grown on plates. Yeast containing ataxin-1[82]/LANP turned blue in a shorter time (15 min) than yeast containing ataxin-1[30]/LANP (~ 4 h), or mouse ataxin-1[2]/LANP (~ 7 h) (Fig. 2a). Second, β -gal activity was quantified in liquid cultures of yeast containing LANP and ataxin-1 with either 2, 30 or 82 glutamines. The interaction of ataxin-1 and LANP was significantly increased (~ 10 -fold) with 82 glutamines compared with 30 or 2 glutamines ($F_{2,10} = 302.45$, $P < 0.0001$; Fig. 2b). These results demonstrate that the mutant form of ataxin-1 associates more strongly with LANP.

In situ hybridization studies revealed that *Lanp* mRNA is enriched in cerebellar Purkinje cells and granule neurons³. The highest expression levels in Purkinje cells are on postnatal day 14 (P14), after which the expression decreases to adult levels. Because the first three postnatal weeks are important for neuronal differ-

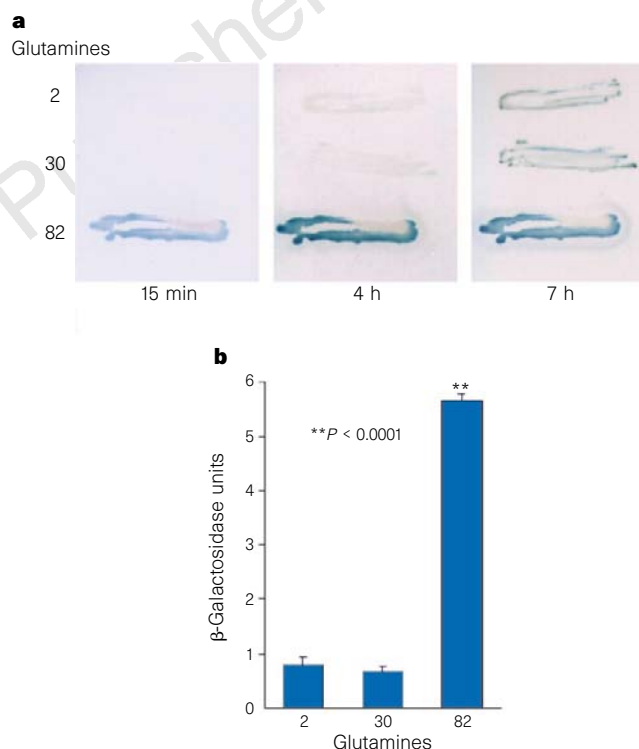


Figure 2 β -Galactosidase assays of the interaction between ataxin-1 and LANP. **a**, Filter-lift assay of β -galactosidase activity in yeast containing LANP and ataxin-1 with either 2, 30 or 82 glutamines. **b**, Liquid assays showing the β -galactosidase activity based on the interaction between LANP and ataxin-1, with either 2, 30, or 82 glutamines in ataxin-1. A significant increase in interaction between ataxin-1 and LANP is noted in the presence of 82 glutamines ($P < 0.0001$). Error bars indicate s.e.m.

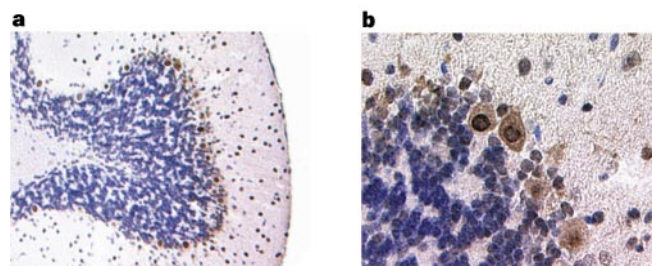


Figure 3 Immunohistochemical staining of LANP in mouse cerebellar cortex. LANP is localized predominantly to the nuclei of Purkinje cells. **a**, Magnification $\times 275$; **b**, $\times 880$.

entiation and synaptogenesis in the rat cerebellum, the temporal expression profile of LANP suggests that this nuclear protein may be involved in cerebellar morphogenesis. We observed a similar temporal expression profile for the *Sca1* gene product. In Purkinje cells, *Sca1* mRNA levels are highest at P14, after which they gradually decrease to adult level¹¹. Immunohistochemical analysis using anti-LANP antisera in murine brain slices showed the highest levels of LANP in the nuclei of cerebellar Purkinje cells, with much weaker signal in the cytoplasm of Purkinje cells and nuclei of granule cells

(Fig. 3). Similarly, ataxin-1 is predominantly nuclear in Purkinje cells, with very low but detectable levels found in the cytoplasm⁴. LANP immunoreactivity was detected, although at lower levels, in the cerebellar deep nuclei, brainstem and thalamus. The expression pattern of LANP in Purkinje cells and the overlapping spatial and temporal expression patterns of both LANP and ataxin-1 make LANP an excellent candidate to be a cell-specific target implicated in SCA1 pathology. Recently, we have identified the human LANP homologue confirming the existence of this protein in humans. The

Figure 4 Colocalization of LANP and ataxin-1 in subnuclear structures by confocal immunofluorescence in COS-7 cells. Anti-Flag antibody (red) and ataxin-1 antisera (green) were used to detect LANP and ataxin-1, respectively. **a**, The distribution of LANP when expressed alone in COS-7 cells is homogeneous in the nucleus. Ataxin-1 with either 30 glutamines (**b**) or 82 glutamines (**c**) shows both a diffuse nuclear staining pattern as well as discrete subnuclear structures in transiently transfected COS-7 cells. COS-7 cells stably expressing ataxin-1 with either 30 (**d-f**), or 92 (**g-i**) glutamines were transiently transfected with LANP. The size of the structures is largest in the cells expressing ataxin-1 with either 82 (**c**) or 92 (**h**) glutamines. **d, g**, Immunofluorescence demonstrates that, in presence of ataxin-1, LANP localizes to the subnuclear structures. **f, i** Overlays demonstrating the colocalization of ataxin-1 and LANP. Magnification, $\times 3,500$.

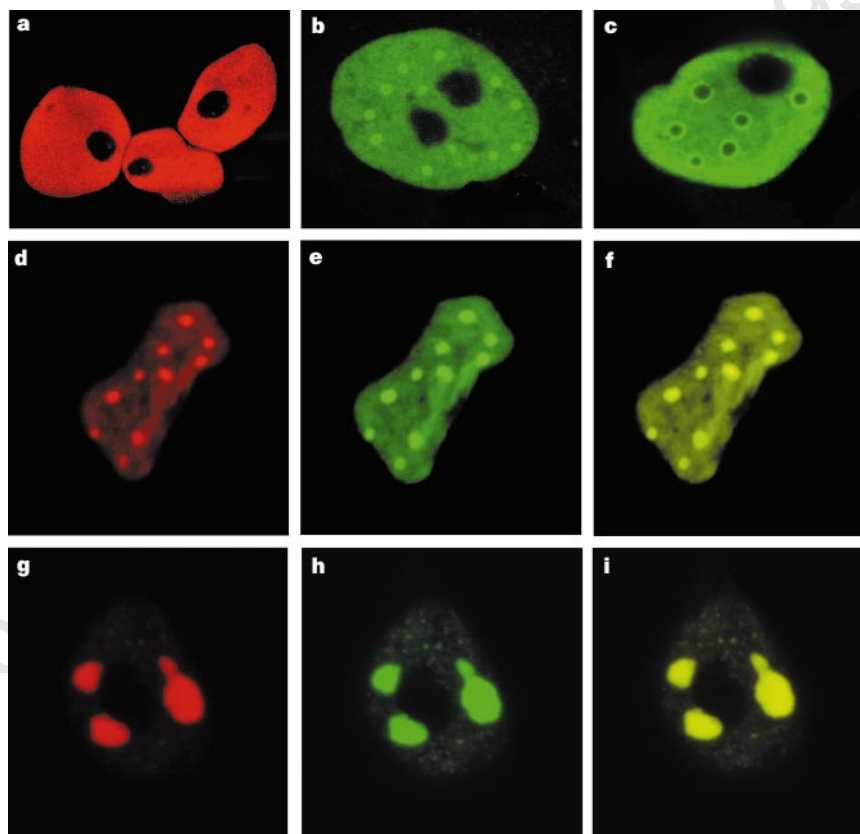
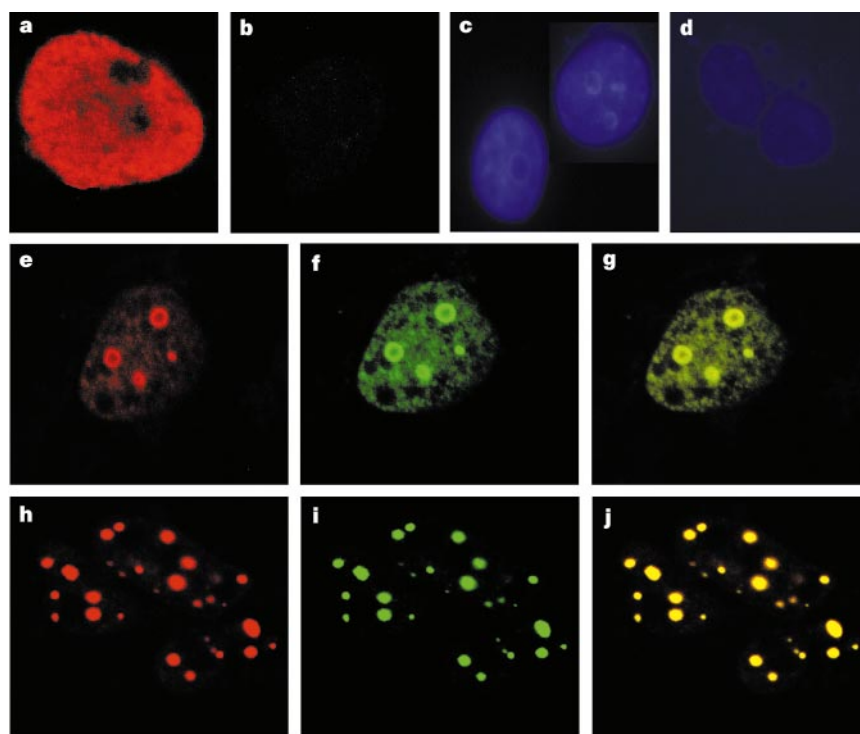


Figure 5 Confocal microscopy of isolated nuclear matrix from COS-7 cells demonstrates localization of LANP to the nuclear matrix in the presence of ataxin-1. Anti-Flag antibody (red) and ataxin-1 antisera (green) were used to detect LANP and ataxin-1, respectively. **a**, Nuclei in unextracted COS-7 cells expressing LANP alone show a homogeneous staining pattern. **b**, Nuclear matrix preparations from LANP-transfected cells show absence of LANP staining. **c, d**, DAPI staining before (**c**) and after (**d**) DNaseI treatment demonstrates the removal of nucleic acids in the nuclear matrix preparation (**d**). **e-j**, Nuclear matrix preparations from COS-7 cells cotransfected with LANP and ataxin-1 containing either 30 (**e-g**) or 82 (**h-j**) glutamines demonstrate that LANP localizes to the nuclear matrix through its interaction with ataxin-1 (**g, j**). Magnification $\times 3,500$.



mouse and human proteins have 89% amino-acid identity (GenBank accession numbers AF022957 and AF025684, respectively).

Ataxin-1 localizes to subnuclear structures in COS-1 cells and cerebellar Purkinje cells of SCA1 patients and transgenic mice¹². The number and the size of these structures depend on the number of glutamines, and ataxin-1 is present in the nuclear matrix fraction¹². We analysed the localization pattern of LANP in COS-7 cells that were either transiently expressing ataxin-1 with 30 or 82 glutamines, or stably expressing ataxin-1 with 30 or 92 glutamines. When LANP was transfected alone into COS-7 cells, a homogeneous staining pattern was noted in the nucleus (Fig. 4a). In contrast, COS-7 cells expressing either LANP and ataxin-1[30] (Fig. 4d–f) or LANP and expanded ataxin-1[82] or ataxin-1[92] (Fig. 4g–i) demonstrated an altered staining pattern of LANP, with LANP colocalizing with ataxin-1 in the subnuclear structures (Fig. 4f, i). These results demonstrate that LANP, which by itself does not form nuclear aggregates, is recruited to the subnuclear structures through its interaction with ataxin-1. To demonstrate the specificity of the interaction between ataxin-1 and LANP in cultured cells, we overexpressed ataxin-1 and the nuclear transcription factor SP1 in COS-7 cells and demonstrated that SP1 does not colocalize with ataxin-1 in the subnuclear structures (data not shown). Finally, we analysed the localization of LANP using an *in situ* nuclear matrix preparation¹³, and observed that when LANP was expressed alone it was removed efficiently from the nucleus after detergent extraction (Fig. 5a, b). In contrast, when LANP was coexpressed with ataxin-1 in COS-7 cells it localized to the salt-resistant subnuclear structures associated with the nuclear matrix (Fig. 5e–j). These results demonstrate that LANP is sequestered and localized to the nuclear matrix through its interaction with ataxin-1. Co-immunoprecipitation experiments failed to demonstrate an interaction between ataxin-1 and LANP. Biochemical isolation of LANP and ataxin-1 using an *in vitro* nuclear matrix extraction protocol^{13,14} demonstrated that LANP was easily removed by detergent treatment (data not shown). These data, together with the *in situ* immunolocalization data, lead us to believe that the structural integrity of the nucleus is essential for the interaction between ataxin-1 and LANP.

The data available to date strongly support a gain-of-function model of pathogenesis for SCA1. Transgenic mice overexpressing a mutant form of a human SCA1 cDNA in Purkinje cells develop ataxia and Purkinje-cell degeneration¹⁵. In contrast, mice lacking the *Scal* gene have learning and memory deficits without evidence of ataxia¹⁶. The specificity of the neuronal phenotype is probably determined both by the expression levels of ataxin-1 and by interactions with neuronal specific nuclear targets. The temporal and cell-specific expression pattern of LANP in cerebellar Purkinje cells, as well as its enhanced interaction with mutant ataxin-1, renders LANP an excellent candidate to be a cell-specific target in SCA1 pathogenesis. The fact that mutant ataxin-1 leads to subnuclear alterations in SCA1 patients, SCA1 transgenic mice, and COS-1 cells¹² suggests that the pathogenesis of SCA1 is mediated through the disturbance of normal nuclear function.

LANP belongs to a group of proteins that contain LRRs. Proteins containing LRRs occur in a wide range of organisms, and several are known to be involved in early morphogenesis and cell adhesion¹⁷. Although very few properties are shared amongst LRR-containing proteins, most are involved in protein–protein interactions, and about half are involved in known signal-transduction pathways^{10,17}. The localization of LANP to the large subnuclear structures formed by mutant ataxin-1 may keep LANP from exerting its normal function, lead to the sequestration of additional nuclear factors, or promote its interaction with incorrect nuclear targets. Although the precise function of LANP is not yet known, its temporal expression pattern during cerebellar development suggests that it may be involved in cerebellar morphogenesis. Generation of mice with a null mutation for LANP or overexpressing LANP in Purkinje

cells should provide further insight into the role of LANP in SCA1 pathology. □

Methods

Plasmids and yeast strains. A human SCA1 cDNA containing 82 CAG repeats was subcloned in the pAS2-CYH2 vector, which contains the GAL4 DNA-binding domain¹⁸, and transformed into the yeast strain Hf7c¹⁹. A Hf7c yeast clone expressing ataxin-1[82] was transformed with a mouse brain library cloned in the pACT2 vector containing the GAL4 activation domain²⁰. Yeast clones growing in SC-His/-Trp/-Leu were tested for β -gal activity using the filter-lift method²¹. To assess the specificity of the interaction between ataxin-1[82] and LANP, Y187 yeast, containing the first 147 amino acids of LANP was mated with Hf7c containing ataxin-1[30], ataxin-1[82], Lamin, SNF1, Tat, Cdc25, p53, Rb, Rev, cyclin D1, Cdk2 and Cdk6, to form diploids. Diploids were tested for β -gal activity. The bait constructs containing ataxin-1 with 30 glutamines and different regions of the SCA1 cDNA were as previously described⁹. A partial murine *Scal* cDNA (amino acids 1–226) containing two CAG repeats was cloned into the pAS2-CYH2 vector.

Liquid assay of β -galactosidase activity. A liquid assay of four independent Y190 transformants containing LANP and ataxin-1 with either 2, 30 or 82 glutamines, each assay performed in triplicate, was performed as described²². The values (mean $\pm$ s.d.) were averaged and ANOVA was performed to evaluate the statistical significance.

Immunohistochemistry and immunofluorescence. Immunohistochemical staining was performed using a bovine anti-LANP polyclonal antibody³ on mouse brain sections⁴. Transient expression of ataxin-1 and LANP in COS-7 cells was done according published protocols¹⁵. Full-length mouse *Lamp* was subcloned in the pFlag-CMV-2 vector (Kodak). COS-7 cells stably expressing human ataxin-1 with either 2, 30 or 92 glutamines were generated using the pcDNA3.1/HIS-C expression vector system (Invitrogen). Cells were transfected and passaged at least three times in the presence of 250 μ g ml⁻¹ G418 in DMEM (Gibco/BRL). Cells were fixed for 10 min in 3.7% formaldehyde in phosphate-buffered saline (PBS) 48 h after transfection, permeabilized in acetone at -20°C , and incubated for 60 min at 37°C in a humid chamber with rabbit polyclonal antibodies (1:100) recognizing ataxin-1 (11750V)⁴ and/or mouse monoclonal antibodies (1:100) M2 (Kodak) recognizing the Flag epitope. Subsequently, cells were incubated with either anti-rabbit-FITC or anti-mouse-TRITC (1:100, Vector Laboratories), mounted in antifade solution (Vectashield mounting medium, Vector Laboratories), and analysed by laser confocal microscopy (BioRad 1024).

Cytoskeletal and nuclear matrix preparation. Cytoskeletal and nuclear matrix preparations were carried out 48 h after transfection as described¹³. Briefly, coverslips were rinsed with PBS, extracted on ice with cytoskeletal buffer contain (in mM) 10 PIPES, pH 6.8, 300 sucrose, 100 NaCl, 3 MgCl₂, 1 EGTA, 0.5% Triton X-100, 1.2 PMSE, 2 vanadylribonucleoside complex and 1 $\times$ protease inhibitors (Complete, Boehringer). For cytoskeletal preparations, cells were fixed with 3.7% formaldehyde for 20 min on ice and rinsed twice with PBS before immunodetection. For the nuclear matrix preparations, cytoskeletal coverslips were digested with 400 U ml⁻¹ of RNase-free DNaseI (Boehringer) in digestion buffer (cytoskeletal buffer with 50 mM NaCl) for 50 min at 32°C followed by extraction with 0.25 M ammonium sulphate and 2 M NaCl in digestion buffer. Cells were fixed with 3.7% formaldehyde in digestion buffer on ice, rinsed in PBS, and incubated with blocking solution (5% Blotto, BioRad) for 30 min at room temperature. Immunodetection was performed as above. Cells were counterstained in mounting media containing DAPI (1 μ g ml⁻¹), and analysed by laser confocal microscopy (BioRad 1024).

Received 9 July; accepted 8 September 1997.

1. Zoghbi, H. & Orr, H. Spinocerebellar ataxia type 1. *Semin. Cell Biol.* **6**, 29–35 (1995).
2. Orr, H. *et al.* Expansion of an unstable trinucleotide (CAG) repeat in spinocerebellar ataxia type 1. *Nature Genet.* **4**, 221–226 (1993).
3. Matsuoka, K. *et al.* A nuclear factor containing the leucine-rich repeats expressed in murine cerebellar neurons. *Proc. Natl Acad. Sci. USA* **91**, 9670–9674 (1994).
4. Servadio, A. *et al.* Expression analysis of the ataxin-1 protein in tissues from normal and spinocerebellar ataxia type 1 individuals. *Nature Genet.* **10**, 94–98 (1995).
5. Fields, S. & Song, O.-K. A novel genetic system to detect protein-protein interactions. *Nature* **340**, 245–246 (1989).
6. Kordeli, E., Lambert, S. & Bennett, V. AnkyrinG: A new ankyrin gene with neural-specific isoforms localized at the axonal initial segment and node of Ranvier. *J. Biol. Chem.* **270**, 2352–2359 (1995).
7. Wong, Y. H., Kalmbach, S. J., Hartman, B. K. & Sherman, W. R. Immunohistochemical staining and enzyme activity measurements show myo-inositol-1-phosphate synthase to be localized in the

- vasculature of brain. *J. Neurochem.* **48**, 1434–1442 (1987).
8. Smith, K. E. *et al.* Molecular cloning of an orphan transporter. A new member of the neurotransmitter transporter family. *FEBS Lett.* **357**, 86–92 (1995).
 9. Koshy, B. *et al.* Spinocerebellar ataxia type-1 and spinobulbar muscular atrophy gene products interact with glyceraldehyde-3-phosphate dehydrogenase. *Hum. Mol. Genet.* **5**, 1311–1318 (1996).
 10. Kobe, B. & Deisenhofer, J. The leucine-rich repeat: a versatile binding motif. *Trends Biochem. Sci.* **19**, 415–421 (1994).
 11. Banfi, S. *et al.* Cloning and developmental expression analysis of the murine homolog of the spinocerebellar ataxia type 1 gene (*Sca1*). *Hum. Mol. Genet.* **5**, 33–40 (1996).
 12. Skinner, P. *et al.* Ataxin-1 with an expanded glutamine tract alters nuclear matrix-associated structures. *Nature* **389**, 971–974 (1997).
 13. Fey, E. G. & Penman, S. Nuclear matrix proteins reflect cell type of origin in cultured human cells. *Proc. Natl Acad. Sci. USA* **85**, 121–125 (1988).
 14. Berezney, R. & Coffey, D. S. Nuclear matrix isolation and characterization of a framework structure from rat liver nuclei. *J. Cell Biol.* **73**, 616–632 (1977).
 15. Burright, E. *et al.* SCA1 transgenic mice: a model for neurodegeneration caused by an expanded CAG trinucleotide repeat. *Cell* **82**, 937–948 (1995).
 16. Matilla, A. *et al.* Ataxin-1, the SCA1 gene product, is required for learning tasks mediated by both the hippocampus and cerebellum. *Proc. Natl Acad. Sci. USA* (submitted).
 17. Kobe, B. & Deisenhofer, J. Proteins with leucine-rich repeats. *Curr. Opin. Struct. Biol.* **5**, 409–416 (1995).
 18. Harper, J. W., Adami, G. R., Wei, N., Keyomarsi, K. & Elledge, S. J. The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. *Cell* **75**, 805–816 (1993).
 19. Feilott, H. E., Hannon, G. J., Ruddell, C. J. & Beach, D. Construction of an improved host strain for two hybrid screening. *Nucleic Acids Res.* **22**, 1502–1503 (1994).
 20. Durfee, T. *et al.* The retinoblastoma protein associates with the protein phosphatase type 1 catalytic subunit. *Genes Dev.* **7**, 555–569 (1993).
 21. Breeden, L. & Nasmyth, K. Regulation of the yeast HO gene. *Cold Spring Harb. Symp. Quant. Biol.* **50**, 643–650 (1985).
 22. Rose, M. & Botstein, D. Construction and use of gene fusions to lacZ (β -galactosidase) that are expressed in yeast. *Methods Enzymol.* **101**, 167–180 (1983).

Acknowledgements. We thank S. Elledge for the Y187 and Y190 yeast strains, pAS2-CYH2 and pACT2 vectors, and constructs expressing Lamin, SNF1, Tat, Cdc25, p53, Rb, Rev, cyclin D1, Cdk2 and Cdk6; E. N. Olson for the mouse brain cDNA library; D. Beach for the H17c yeast strain; D. Armstrong for neuropathological expertise; S. Y. Tsai for the expression vector containing SP1; R. C. Atkinson for assistance with confocal microscopy; R. Phillips and B. Antalffy for technical expertise; and A. Beaudet for critically reading the manuscript. This work was supported by grants from the NIH to H.Y.Z. and H.T.O. and by the Baylor College of Medicine Mental Retardation Research Center. A.M. was supported by a postdoctoral fellowship from the Spanish Ministerio de Educación y Ciencia.

Correspondence and requests for materials should be addressed to H.Y.Z. (e-mail: hzoghbi@bcm.tmc.edu).

Cutaneous lymphocyte antigen is a specialized form of PSGL-1 expressed on skin-homing T cells

Robert C. Fuhlbrigge*, J. David Kieffer*, Dieter Armerding* & Thomas S. Kupper

Harvard Skin Disease Research Center, Division of Dermatology, Department of Medicine, Brigham and Women's Hospital, 75 Francis Street, Boston, Massachusetts 02115, USA

* These authors contributed equally to this work.

T cells play a pathogenic role in many inflammatory and certain malignant skin diseases, including psoriasis, atopic and allergic contact dermatitis, and cutaneous T-cell lymphoma^{1–6}. Memory T cells that infiltrate the skin express a unique skin-homing receptor called cutaneous lymphocyte-associated antigen (CLA), a carbohydrate epitope that facilitates the targeting of T cells to inflamed skin^{1,7,8}. CLA is defined by both its reactivity with a unique monoclonal antibody, HECA-452, and its activity as a ligand for E-selectin^{2,9–11}, but the structure of the protein component of CLA has not previously been defined. Here we report that CLA is an inducible carbohydrate modification of P-selectin glycoprotein ligand-1 (PSGL-1), a known surface glycoprotein that is expressed constitutively on all human peripheral-blood T cells. Cultured peripheral-blood T cells can be differentiated into CLA-bearing cells, which bind both E-selectin and P-selectin, or CLA-negative cells, which bind P-selectin but do not bind E-selectin, suggesting that there is independent regulation of selectin-binding phenotypes. We propose that differential post-translational modification of a single cell-surface receptor, PSGL-1, mediated by fucosyltransferase VII, serves as a mechanism for regulating tissue-specific homing of memory T cells.

The accumulation of memory T cells in non-lymphoid peripheral tissues is typically initiated by the interaction of glycoprotein ligands on the T-cell surface with E-selectin or P-selectin expressed on vascular endothelium^{12–15}. Previous studies have shown that distinct subsets of memory T cells tether and roll on E-selectin and P-selectin *in vitro*, suggesting that T-cell binding to E-selectin and P-selectin is mediated by independent cell-surface receptors^{2,11,16}. T cells are known to bind P-selectin through PSGL-1, a well-characterized mucin-like glycoprotein that seems to be the sole physiological ligand for P-selectin¹⁷. To function as an effective P-selectin ligand, the PSGL-1 protein must be modified by appropriate glycosylation and tyrosine sulphation. The E-selectin ligands on T cells, in contrast, are not as well characterized. Although CLA is expressed by a sizeable minority of peripheral-blood T cells, activation and expansion under standard T-cell culture conditions results in loss of expression, complicating the analysis of its structure and function. Methods have been described that partly reverse this inhibition, but these result in only low-level expression^{8,18}.

We have developed conditions, using a serum-free medium (XVIVO15), that result in the induction and maintenance of high levels of CLA on virtually all peripheral blood T cells (Fig. 1b). These conditions are similar to those reported to induce CLA expression on cultured vaccine-primed lymph-node cells¹⁹. Parallel activation and culture in conventional T-cell growth media with fetal calf serum (CG/RPMI-FCS) resulted in an initially low stimulation (not shown), followed by loss of CLA expression within 7 days (Fig. 1c). The basis for these properties of serum-free medium is yet to be established, although the addition of FCS to XVIVO15 medium only partly inhibited CLA expression (not shown).

To confirm that CLA expressed on T cells cultured *in vitro* was similar to that on unstimulated peripheral blood T cells, HECA-452-reactive cells were immunomagnetically purified from peripheral-blood mononuclear cells and compared with T cells expanded *in vitro*. The level of CLA expressed on HECA-selected

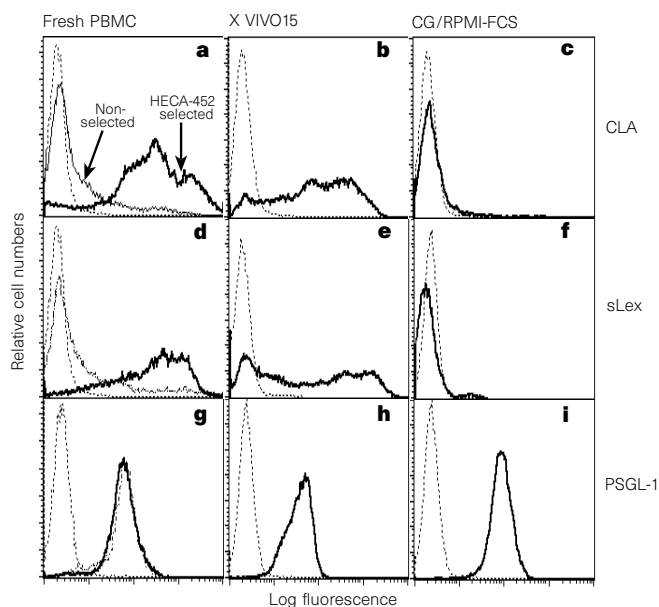


Figure 1 Expression of CLA, sLex^x and PSGL-1 by fresh and cultured PBMCs. Freshly prepared human PBMCs were stained unselected and after selection for HECA-452-expressing cells. HECA-selected cells (thick line) expressed high levels of CLA (HECA-452 mAb) (a) and sLex^x (CSLEX-1 mAb) (d) relative to unselected cells (thin line). PBMCs cultured in XVIVO15 medium expressed high levels of CLA (b) and sLex^x (e); PBMCs cultured in CG/RPMI-FCS medium showed minimal staining for either CLA (c) or sLex^x (f). PSGL-1 was expressed at similar levels in all groups (g–i). The broken lines represent staining with isotype-matched control antibodies.

cells was similar to that observed on cells cultured in XVIVO15 medium (Fig. 1a). Expression of sLe^x, defined by reactivity with the monoclonal antibody CSLEX-1, was similar to HECA-452 expression (Fig. 1d–f). In contrast, cell-surface expression of PSGL-1 was similar on all T cells, and was not significantly affected by culture conditions (Fig. 1g–i).

Selectin ligand function was characterized in a parallel-plate flow chamber under physiological shear flow conditions. CLA-positive peripheral-blood T cells, activated and cultured in XVIVO15 medium, tethered in shear flow and rolled on E-selectin and P-selectin with high efficiency (Fig. 2a–c). In contrast, CLA-negative T cells, maintained in conventional media, did not form stable tethers to E-selectin, even under static conditions, which is consistent with their lacking the CLA/HECA-452 epitope (Fig. 2a, b). There was no significant difference in P-selectin binding activity between CLA-positive and negative cultured T cells (Fig. 2a, c). In these qualitative studies, fresh peripheral-blood mononuclear cells selected for CLA expression showed higher tethering rates to E-selectin and P-selectin than XVIVO15-cultured cells (Fig. 2d), which is consistent

with the higher level of CLA expressed on these cells (Fig. 1). Resistance to detachment was quite similar, however, suggesting that the intrinsic binding properties of the selectin ligands on fresh and cultured cells are the same (Fig. 2e, f). However, true quantitative comparisons remain to be completed. The observation that CLA-positive and negative cell populations tethered and rolled on P-selectin with comparable efficiency confirms and extends previous observations^{11,16,20} and indicates that neither CLA nor sLe^x is required for P-selectin ligand activity on T cells, in contrast to the leukocytes. The addition of EDTA resulted in the complete abrogation of P-selectin binding (not shown), confirming that this is a calcium-dependent interaction.

We next sought to characterize the protein structure of CLA. Immunoblotting of XVIVO15-cultured T-cell lysates with the monoclonal antibody HECA-452 revealed a highly reproducible pattern of immunoreactivity. Two discrete and specific bands, migrating with apparent relative molecular mass (M_r) of 140K and 240K under standard reducing conditions, were detected in CLA-positive cell populations (Fig. 3, lanes 1 and 3). In contrast, no specific HECA-452 immunoreactivity could be demonstrated in CLA-negative T-cell lysates (cells grown in CG/RPMI-FCS) (Fig. 3, lane 2). CLA-positive cells purified from unstimulated normal human peripheral blood showed a pattern of immunoreactivity that was indistinguishable from that of XVIVO15-cultured CLA-positive T cells (Fig. 3, lanes 3, 4). Similar bands were also

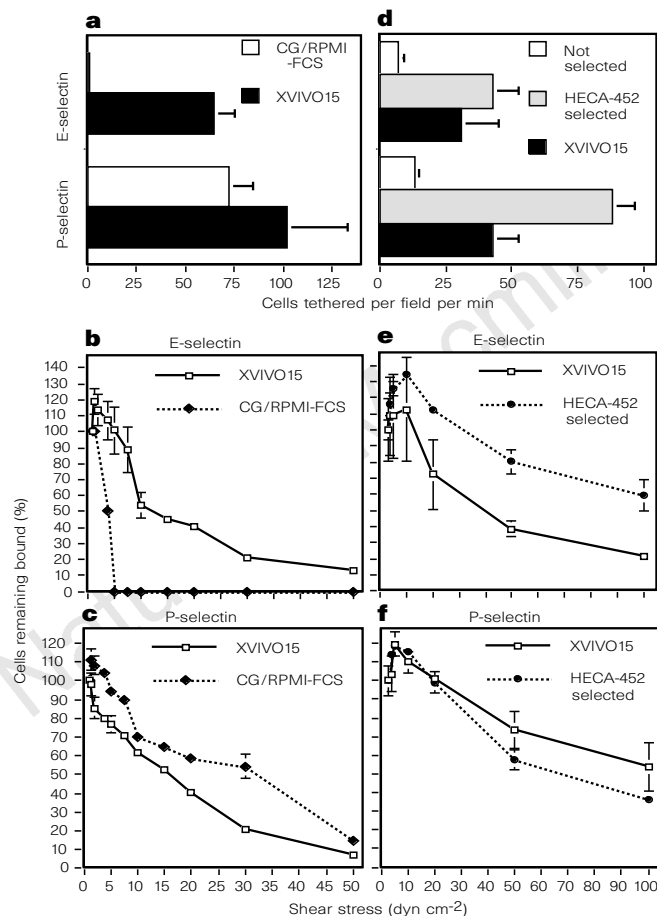


Figure 2 Tethering and rolling of fresh and cultured PBMCs on selectins in physiological shear flow. **a**, PBMCs cultured in XVIVO15 bound to E-selectin with high efficiency, whereas cells cultured in CG/RPMI-FCS medium formed few stable tethers. **b**, XVIVO cells rolling on E-selectin displayed high resistance to detachment by shear stress, whereas the few CG/RPMI-FCS cells observed that bound to E-selectin detached at low shear stress. In contrast, both cell populations bound efficiently to P-selectin (**a**) and displayed nearly equivalent resistance to detachment (**c**). Fresh PBMCs selected for HECA-452 expression tethered with high efficiency to both E-selectin and P-selectin (**d**). Unselected cells showed fewer tethers, consistent with expression of CLA on only a fraction of the cells. HECA-selected cells showed resistance to detachment by shear stress equivalent to XVIVO15 cultured cells on both E-selectin (**e**) and P-selectin (**f**).

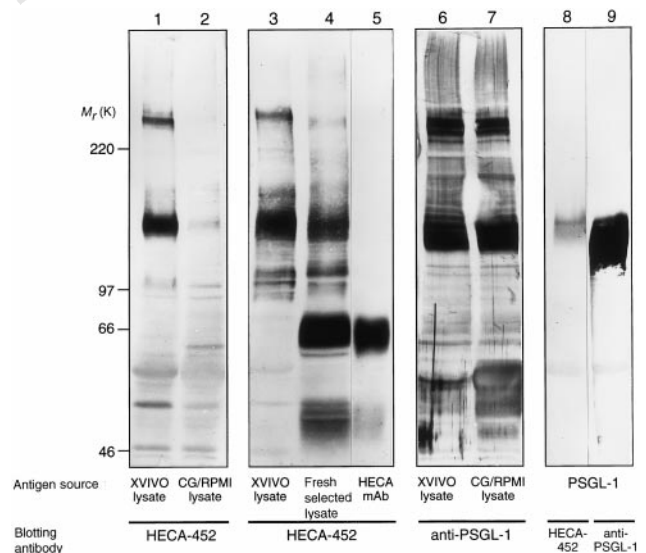


Figure 3 Immunoblotting of peripheral-blood T-cell lysates. Lysates prepared from fresh or cultured PBMCs were subjected to SDS-PAGE and immunoblotted with HECA-452 and anti-PSGL-1 antibodies as indicated. The mAb HECA-452 consistently and specifically identified two bands of apparent M_r 240K and 140K in cells cultured in XVIVO15 (lane 1) but did not identify specific bands in cells cultured in CG/RPMI-FCS (lane 2), consistent with the pattern of CLA expression by FACS (Fig. 1). Fresh PBMCs selected for HECA-452 expression (lane 4) contained the same two bands as XVIVO15-cultured cells (lane 3) when blotted with HECA-452. The apparent difference in mobility of the upper bands of lanes 3 and 4 is not evident in similar blots. The broad band near 50K in lysates of freshly selected PBMCs is the HECA-452 antibody used for selection that was recognized by the anti-immunoglobulin developing reagent. HECA-452 antibody alone was run in a separate lane to confirm this relationship (lane 5). Immunoblotting with anti-PSGL-1 antiserum identified the same 240K and 140K bands in both XVIVO15 and CG/RPMI-FCS cultured cell lysates (lanes 6 and 7, respectively). To confirm that HECA-452 could recognize PSGL-1 glycoprotein, recombinant soluble PSGL-1 produced in COS cells co-transfected with fucosyltransferase was immunoblotted with HECA-452 (lane 8) and anti-PSGL-1 (lane 9). A 140K glycoprotein representing the monomeric form of PSGL-1 was identified by both antibodies.

observed in lysates of a CLA-positive T-cell clone (not shown).

To compare CLA with a selectin ligand of known structure, we performed immunoblotting with an anti-PSGL-1 antibody. PSGL-1 is a glycoprotein of M_r 110K–120K that is expressed as a cell-surface homodimer of 220K–240K (refs 21–23). To our surprise, anti-PSGL-1 antiserum identified two bands (Fig. 3, lanes 6 and 7) that co-migrated with HECA-452-reactive bands in parallel blots (Fig. 3, lanes 1 and 2), with two notable differences. First, the PSGL-1 bands were broader than the corresponding HECA-452 bands. Second, and more strikingly, the intensity of the PSGL-1 bands was equivalent in both CLA-positive (XVIVO15 medium) and CLA-negative (CG/RPMI-FCS medium) T-cell lysates. Identical results were obtained with a monoclonal anti-PSGL-1 antibody (not shown). There are at least two plausible interpretations of these results. CLA and PSGL-1 may be structurally unrelated but have similar apparent relative molecular masses. However, a more intriguing hypothesis is that CLA and PSGL-1 share a common protein backbone, expressed on all T cells, that can be differentially decorated with carbohydrate epitopes to generate CLA on a subset of T cells. The hypothesis that CLA is a specially glycosylated isoform of PSGL-1 was tested through several independent approaches.

To confirm a relationship between CLA and PSGL-1 on material from a separate source, we performed blotting studies using soluble monomeric recombinant PSGL-1 produced in cells co-transfected

with PSGL-1 and $\alpha(1,3/1,4)$ -fucosyltransferase III (ref. 23). This product exhibits both P-selectin and E-selectin binding activity. Monomeric PSGL-1 reacted with both HECA-452 and anti-PSGL-1 antibodies (Fig. 3, lanes 8 and 9) and migrated in a manner indistinguishable from that of the 140K band from CLA-positive T-cell lysates. As noted with T-cell lysates, HECA-452 bound only a restricted portion of the total anti-PSGL-1 reactive material, suggesting that only a subset of the PSGL-1 molecules were decorated with the HECA-452 carbohydrate epitope.

Although PSGL-1 is a disulphide-linked homodimer and would be expected to migrate as a single band under conventional reducing conditions, our experiments consistently revealed two bands in both HECA-452 and anti-PSGL-1 blots. Previous reports have indicated that PSGL-1 is relatively resistant to complete reduction²¹. In support of the idea that the 240K species is a dimer and the 140K protein a monomer, treatment of lysates with 5% 2-mercaptoethanol for 72 h at 4°C resulted in complete conversion of both the HECA-452 reactive and anti-PSGL-1 reactive material to single bands migrating with an apparent M_r of 140K (Fig. 4). Identical results were obtained from lysates of freshly isolated CLA-positive cells (not shown). The fully reduced 140K band from fresh and cultured cells was indistinguishable from the monomeric recombinant PSGL-1 shown in Fig. 3, lending further support to the hypothesis that CLA and PSGL-1 are identical.

To determine whether the epitopes recognized by the respective specific antibodies were co-expressed on a single glycoprotein, anti-PSGL-1 immunoprecipitates of CLA-positive T cells were subjected to SDS-PAGE and immunoblotted with HECA-452 antibody (Fig. 5). HECA-452 reacted with the immunoprecipitated PSGL-1, indicating that the CLA epitope is expressed directly on PSGL-1 glycoprotein isolated from T cells. This confirms that these antigens share a common protein core. Preliminary electron microscopy studies confirm that T-cell CLA expression is limited to microvilli (not shown), consistent with its identity as PSGL-1, and argue against expression on other surface elements such as glycolipid.

Our data demonstrate that CLA, which defines the skin-homing T-cell population in humans, is expressed on T cells as an epitope of a single cell-surface protein, PSGL-1. This observation, in addition to our previous data^{2,11}, indicates that PSGL-1 is the principle glycoprotein responsible for the interaction of memory T cells with both E-selectin and P-selectin in peripheral tissues. *In vivo*, CLA is thought to be induced on T cells undergoing naive to memory transition in lymph nodes that drain the skin⁸. We propose that CLA is produced by post-translational glycosylation of PSGL-1, which is expressed constitutively, through the activity of FucTVII, an $\alpha(1,3)$ -fucosyltransferase found in leukocytes that bind E-selectin^{20,24}. In support of this hypothesis, transfection of Jurkat T cells with FucTVII generates CLA-positive cell lines that bind both E-selectin and P-selectin¹⁹ (R.C.F., unpublished observation), and CLA-positive cultured human T cells contain FucTVII protein whereas CLA-negative T cells do not (J.D.K., unpublished observation). The observation that CLA-negative T cells bind P-selectin, but not E-selectin, implies separate modification pathways regulating expression of E-selectin and P-selectin ligand activity. The ability differentially to regulate the homing of memory T cells to skin through the modification of an ubiquitous cell-surface protein adds an important dimension to our understanding of T-cell homing. Identification of the variables controlling the expression and persistence of these glycosylation and homing phenotypes may provide attractive new targets for therapy of T-cell-mediated inflammatory diseases. □

Methods

Cell preparations and tissue culture. Peripheral-blood mononuclear cells (PBMCs) were prepared by density gradient separation (Ficoll-Hypaque 1.077; Sigma) of peripheral blood or cells collected during platelet pheresis of normal donors. PBMCs were analysed directly (unselected) or selected by incubation with the monoclonal antibody (mAb) HECA-452 and anti-rat immunoglobulin

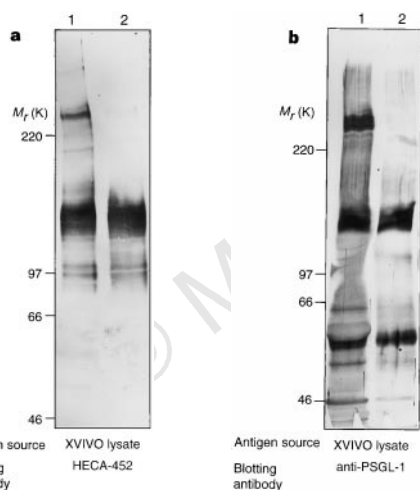


Figure 4 Immunoblot of lysates exposed to enhanced reducing conditions. Lysates of PBMCs cultured in XVIVO15 medium were used untreated (lane 1) or treated with 5% 2-mercaptoethanol for 72 h at 4°C (lane 2) before electrophoresis under standard reducing conditions and immunoblotting with HECA-452 mAb (a) or anti-PSGL-1 antiserum (b). The 240K band, but not the 140K band, identified by both antibodies was eliminated by enhanced reduction.

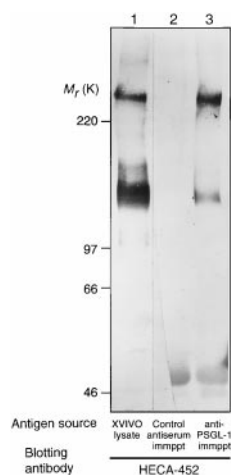


Figure 5 Immunoblot of immunoprecipitated PSGL-1 with HECA-452. Lysate from PBMCs cultured in XVIVO15 was immunoprecipitated with normal rabbit serum or anti-PSGL-1 and immunoblotted with HECA-452 mAb. The 240K and 140K bands recognized by HECA-452 in the original lysate (lane 1) were evident in the anti-PSGL-1 immunoprecipitate (immpt) (lane 3) but not in control rabbit antiserum immunoprecipitate (lane 2), indicating specific immunoprecipitation of the CLA molecule by anti-PSGL-1 antiserum. The relative enhancement of the upper band relative to the lower band is highly reproducible. This probability reflects an advantage of oligomeric PSGL-1 over monomeric PSGL-1 for incorporation into immune complexes, and not a change in HECA-452 reactivity.

magnetic beads (MACS; Miltenyi Biotec) according to the manufacturer's instructions. Selected cells used in these studies were typically >90% CD3⁺ and <5% CD16⁺ (not shown). Cultured cells were generated by resuspending PBMCs in XIVO15 medium (Bio-Whittaker) or a 1:1 mixture of CG medium (Vitromex, Germany) and RPMI 1640 medium (Mediatech) supplemented with 10% fetal calf serum (CG/RPMI-FCS). Cells were initially incubated in 24-well plastic tissue-culture plates (Costar) coated with anti-human CD3 mAb (Immunotech) on goat anti-mouse immunoglobulin (Southern Biotechnology Associates). Subsequent passage (every 3–4 days) was performed in uncoated plates using the same media supplemented with recombinant human interleukin-2 (100 U ml⁻¹; Amgen). Cells were analysed between 7 and 14 days after stimulation. Cultured cells showed undetectable levels of CD16 or CD19 staining.

FACS analysis. For antibody staining, 5 × 10⁵ cells were incubated with 2–5 µg mAb HECA-452 (rat IgM mAb to CLA) (L. Picker)¹, CSLEX-1 (mouse IgG, ATCC HB8580)²⁵, or either rabbit anti-PSGL-1 antiserum or anti-PSGL-1 mAb (PSL-275, mouse IgG1 mAb)²³ for 1 h at 4 °C, washed and incubated with appropriate FITC-conjugated goat anti-immunoglobulin antisera (Zymed Laboratories or Southern Biotechnology Associates) at 1:20 to 1:50 dilution for 30 min at 4 °C. Flow cytometry was performed on a FACScan IV (Becton Dickinson) using CellQuest software (version 1.2). Results are representative of multiple independent observations for each data set.

Flow analysis. Cells were analysed in a parallel-plate flow chamber using substrates coated with protein A and E-selectin or P-selectin IgG chimaera as described²⁶. Washed cells were resuspended in Hank's balanced salt solution with 2 mM CaCl₂ at 0.5–1 × 10⁶ cells ml⁻¹ and immediately infused into the chamber. Tethering was observed at 1 dyn cm⁻² for 1 min, followed by stepwise increases in wall shear stress every 15 s to 100 dyn cm⁻². More than 95% of attached cells rolled in all experiments. Calcium-dependent binding was confirmed by perfusion of 5 mM EDTA with >95% release of bound cells for all samples studied. All experiments were observed in real time and videotaped for analysis. Multiple experiments were performed for each data set, with representative results shown here. Each data point represents the mean and range of two independent observations.

Immunoblotting. Cell lysates were generated as described¹, except that lysates contained 2 × 10⁸ cells ml⁻¹ and were incubated for 2 h at 4 °C before the 10,000g spin. Recombinant soluble PSGL-1 is an engineered product that lacks the normal cytoplasmic tail of PSGL-1 and does not form stable dimers in solution²³. General methods for SDS–PAGE under reducing conditions and immunoblotting were as described²⁷, except that proteins were analysed on 7.5% acrylamide (National Diagnostics) gels with half of the conventional amount of crosslinker (0.1% bisacrylamide). In all experiments, the dye front was retained on the gel. Gels were blotted onto Immobilon-P (Millipore), which was subsequently blocked with newborn-calf serum (Mediatech) for 1 h and probed with HECA-452 (2 µg ml⁻¹) or anti-PSGL-1 antisera (1:200) as indicated. Isotype control blots (rat IgM from Zymed; normal rabbit serum from Pel-Freez) were performed in parallel for each condition and did not identify 240K or 140K species in any samples tested. Bound antibody was visualized with species-specific alkaline-phosphatase-coupled anti-immunoglobulin (Zymed) and Western blue alkaline phosphatase substrate (Promega) according to the manufacturer's instructions. For enhanced reduction, lysates were brought to 5% 2-mercaptoethanol and incubated for 72 h at 4 °C before SDS–PAGE. All subsequent steps were identical and performed on treated and untreated lysates in parallel. For immunoprecipitation studies, a fraction with high *M_r* was prepared from PBMC lysates by dilution with PBS and concentration on a Centricon 100 ultrafiltration device. This fraction was incubated with control or specific antisera and precipitated with protein A–agarose (Sigma). Immunoprecipitates were subjected to SDS–PAGE and immunoblotted with HECA-452 mAb as described above.

Received 4 June; accepted 11 September 1997.

1. Picker, L. J., Michie, S. A., Rott, L. S. & Butcher, E. C. A unique phenotype of skin-associated lymphocytes in human. Preferential expression of the HECA-452 epitope by benign and malignant T cells at cutaneous sites. *Am. J. Pathol.* **136**, 1053–1068 (1990).
2. Rossiter, H. et al. Skin disease-related T cells bind to endothelial selectins: expression of cutaneous lymphocyte antigen (CLA) predicts E-selectin but not P-selectin binding. *Eur. J. Immunol.* **24**, 205–210 (1994).
3. Picker, L. J. et al. Differential expression of lymphocyte homing receptors by human memory/effector T cells in pulmonary versus cutaneous immune effector sites. *Eur. J. Immunol.* **24**, 1269–1277 (1994).
4. Abernathy-Carver, K. J., Sampson, H. A., Picker, L. J. & Leung, D. Y. Milk-induced eczema is

associated with the expansion of T cell expression of cutaneous lymphocyte antigen. *J. Clin. Invest.* **95**, 913–918 (1995).

5. Pitzalis, C. et al. Cutaneous lymphocyte antigen positive T lymphocytes preferentially migrate to skin but not to the joint in psoriatic arthritis. *Arthritis Rheum.* **39**, 137–145 (1996).
6. Rook, A. H. & Heald, P. The immunopathogenesis of cutaneous T-cell lymphoma. *Hematol. Oncol. Clin. N. Am.* **9**, 997–1010 (1995).
7. Berg, E. L., Robinson, M. K., Mansson, O., Butcher, E. C. & Magnani, J. L. A carbohydrate domain common to both sialyl Le(a) and sialyl Le(x) is recognized by the endothelial cell leukocyte adhesion molecule ELAM-1. *J. Biol. Chem.* **266**, 14869–14872 (1991).
8. Picker, L. J. et al. Control of lymphocyte recirculation in man. II. Differential regulation of the cutaneous lymphocyte-associated antigen, a tissue-selective homing receptor for skin-homing T cells. *J. Immunol.* **150**, 1122–1136 (1993).
9. Picker, L. J., Kishimoto, T. K., Smith, C. W., Warnock, R. A. & Butcher, E. C. ELAM-1 is an adhesion molecule for skin-homing T cells. *Nature* **349**, 796–799 (1991).
10. Berg, E. L. et al. The cutaneous lymphocyte antigen is a skin lymphocyte homing receptor for the cutaneous lymphocyte-associated antigen, a tissue-selective homing receptor for skin-homing T cells. *J. Exp. Med.* **174**, 1461–1466 (1991).
11. Alon, R., Rossiter, H., Wang, X., Springer, T. A. & Kupper, T. S. Distinct cell surface ligands mediate T lymphocyte attachment and rolling on P- and E-selectin under physiological flow. *J. Cell Biol.* **127**, 1485–1495 (1994).
12. Butcher, E. C. & Picker, L. J. Lymphocyte homing and homeostasis. *Science* **272**, 60–66 (1996).
13. Nolte, D. et al. Leukocyte rolling in venules of striated muscle and skin is mediated by P-selectin, not by L-selectin. *Am. J. Physiol.* **267**, H1637–H1642 (1994).
14. Jones, D. A., McIntire, L. V., Smith, C. W. & Picker, L. J. A two-step adhesion cascade for T cell/endothelial cell interactions under flow conditions. *J. Clin. Invest.* **94**, 2443–2450 (1994).
15. Santamaria-Babi, L. F. et al. Migration of skin-homing T cells across cytokine-activated human endothelial cell layers involves interaction of the cutaneous lymphocyte-associated antigen (CLA), the very late antigen-4 (VLA-4), and the lymphocyte function-associated antigen-1 (LFA-1). *J. Immunol.* **154**, 1543–1550 (1995).
16. Diacovo, T. G. et al. Interactions of human alpha/beta and gamma/delta T lymphocyte subsets in shear flow with E- and P-selectin. *J. Exp. Med.* **183**, 1193–1203 (1996).
17. Kansas, G. S. Selectins and their ligands: current concepts and controversies. *Blood* **88**, 3259–3287 (1996).
18. Leung, D. Y. et al. Bacterial superantigens induce T-cell expression of the skin-selective homing receptor, the cutaneous lymphocyte-associated antigen, via stimulation of interleukin 12 production. *J. Exp. Med.* **181**, 747–753 (1995).
19. Knibbl, R. N. et al. The fucosyltransferase FucT-VII regulates E-selectin ligand synthesis in human T cells. *J. Cell Biol.* **133**, 911–920 (1996).
20. Wagers, A. J., Lowe, J. B. & Kansas, G. S. An important role for the 1, 3 fucosyltransferase, FucT-VII, in leukocyte adhesion to E-selectin. *Blood* **88**, 2125–2132 (1996).
21. Moore, K. L. et al. Identification of a specific glycoprotein ligand for P-selectin (CD62) on myeloid cells. *J. Cell Biol.* **118**, 445–456 (1992).
22. Norgard, K. E. et al. Characterization of a specific ligand for P-selectin on myeloid cells: A minor glycoprotein with sialylated O-linked oligosaccharides. *J. Biol. Chem.* **268**, 12764–12774 (1993).
23. Sako, D. et al. Expression cloning of a functional glycoprotein ligand for P-selectin. *Cell* **75**, 1179–1186 (1993).
24. Lowe, J. B. In *Frontiers in Molecular Biology* (ed. Fukuda, M.) 163–205 (Oxford Univ. Press, 1994).
25. Fukushima, K. et al. Characterisation of sialosylated Lewis x as a new tumour-associated antigen. *Cancer Res.* **44**, 5279–5285 (1984).
26. Fuhlbrigge, R. C., Alon, R., Puri, K. D., Lowe, J. B. & Springer, T. A. Sialylated, fucosylated ligands for L-selectin expressed on leukocytes mediate tethering and rolling adhesions in physiologic flow conditions. *J. Cell Biol.* **135**, 837–848 (1996).
27. Kieffer, J. D., Plopper, G., Ingber, D. E., Hartwig, J. H. & Kupper, T. S. Direct binding of F actin to the cytoplasmic domain of the α2 integrin chain *in vitro*. *Biochem. Biophys. Res. Commun.* **217**, 466–474 (1995).

Acknowledgements. We thank L. Picker for the HECA-452 antibody, and G. Larsen for the soluble recombinant PSGL-1 and antibodies to PSGL-1. We acknowledge the NIAMS and NIAID, NIH for support of this work.

Correspondence and requests for materials should be directed to T.S.K. (e-mail: tskupper@bics.bwh.harvard.edu).

Macrophage-tropic HIV and SIV envelope proteins induce a signal through the CCR5 chemokine receptor

Drew Weissman^{*†}, Ronald L. Rabin[†], James Arthos, Andrea Rubbert, Mark Dybul, Ruth Swofford, Sundararajan Venkatesan, Joshua M. Farber & Anthony S. Fauci

Laboratory of Immunoregulation, Laboratory of Clinical Investigation, Flow Cytometry Unit, and Laboratory of Molecular Microbiology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA

[†] These authors contributed equally to this work.

Human immunodeficiency virus (HIV) and simian immunodeficiency virus (SIV) enter target cells by forming a complex between the viral envelope protein and two cell-surface membrane receptors: CD4 and a 7-span transmembrane chemokine receptor

^{*} Present address: Division of Infectious Diseases, University of Pennsylvania Medical Center, 536 Johnson Pavilion, Philadelphia, Pennsylvania 19104, USA.

(reviewed in refs 1–3). Isolates of HIV that differ in cellular tropism use different subsets of chemokine receptors as entry cofactors: macrophage-tropic HIVs primarily use CCR5, whereas T-cell-tropic and dual-tropic isolates use CXCR4 (refs 1–3) receptors. HIV-mediated signal transduction through CCR5 is not required for efficient fusion and entry of HIV *in vitro*^{4,5}. Here we show that recombinant envelope proteins from macrophage-tropic HIV and SIV induce a signal through CCR5 on CD4⁺ T cells and that envelope-mediated signal transduction through CCR5 induces chemotaxis of T cells. This chemotactic response may contribute to the pathogenesis of HIV *in vivo* by chemo-attracting activated CD4⁺ cells to sites of viral replication^{1,2}. HIV-mediated signalling through CCR5 may also enhance viral replication *in vivo* by increasing the activation state of target cells. Alternatively, envelope-mediated CCR5 signal transduction may influence viral-associated cytopathicity or apoptosis.

The B10 tumour-infiltrating lymphocyte (TIL) cell line is CD4⁺ and expresses, as judged by northern blot analysis, CCR1, CCR2b, CCR5 and STRL33 (refs 6–8). Recombinant transmembrane-deleted forms of gp120/41 (hereafter referred to as gp160) from the tissue-culture-adapted HIV-1 strain NL4-3 (T-cell tropic, or T-tropic), the HIV-1 macrophage-tropic (M-tropic) isolated JR-FL, and the non-syncytia-inducing isolates RW020-9, Th014 and CCA were assayed for their ability to induce calcium flux. Transmembrane-deleted forms of gp120/41 (gp160) and gp120 derived from the SIV M-tropic viral clone PBj1.9 and a gp120 protein derived from the SIV T-tropic clone Mac239 were also used (Table 1). B10 cells stimulated with gp160 derived from the JR-FL M-tropic strain exhibited an increase in intracellular calcium similar to that observed for MIP-1 β (Fig. 1a); the signal was sensitive to pertussis-toxin treatment (data not shown), indicating binding to and signalling through G_i-protein-coupled receptors. To determine the receptor responsible for signal transduction, we did desensitization studies. B10 cells were stimulated with JR-FL gp160 and a calcium flux was observed (Fig. 1b); subsequent exposure to either MIP-1 β (Fig. 1b) or JR-FL gp160 (data not shown) did not result in a calcium flux, suggesting that JR-FL gp160 shares the same receptor with MIP-1 β and that this receptor had been desensitized as a result of the prior stimulus from JR-FL gp160. In contrast, a third exposure to either MIP-1 α (Fig. 1b) or RANTES (data not shown) gave normal calcium flux. These results are consistent with the restricted specificity of MIP-1 β and JR-FL for CCR5 and the broadened specificity of MIP-1 α and RANTES for receptors other than CCR5 (ref. 1). Similarly, treatment with MIP-1 β

inhibited signalling upon subsequent exposure to JR-FL gp160 (Fig. 1c) but had no detectable effect on MIP-1 α -induced signalling (data not shown).

MIP-1 β stimulation of 5- to 18-day-old macrophage-depleted anti-CD3 plus interleukin (IL)-2 stimulated peripheral blood mononuclear cells (PBMC) (peripheral blood leukocytes, or PBL) resulted in significant calcium flux (up to 70% of T cells) in both the CD4⁺ and CD8⁺ populations; however, addition of JR-FL gp160 to PBL resulted in calcium flux only in CD4⁺ T cells (Fig. 2a). Saturating amounts of MIP-1 β inhibited the subsequent response to JR-FL gp160 (Fig. 2b), whereas JR-FL gp160 desensitized the subsequent response to MIP-1 β in the CD4⁺ T cells (Fig. 2c). Three additional recombinant HIV envelopes that use CCR5 as a fusion cofactor were assayed. These three envelopes also produced a calcium flux in B10 cells and anti-CD3 plus IL-2-stimulated CD4⁺ T cells (Table 1), indicating that envelope-mediated signal transduction through CCR5 is not restricted to a single isolate of HIV-1.

To determine the role of the CD4 receptor in gp160-mediated signalling through CCR5, monoclonal antibodies against CD4 were used. An antibody that blocks HIV envelope binding and infection (Leu-3a) diminished JR-FL gp160-induced signalling through CCR5. This antibody had no effect on MIP-1 β signalling (Table 2). In contrast, an antibody that does not compete with envelope for CD4 binding (Leu-3) had no effect on JR-FL gp160-induced signalling through CCR5 (Table 2). In addition, JR-FL gp160 was pretreated with soluble CD4 (sCD4) before addition to PBL. The sCD4/JR-FL gp160 complex, like JR-FL gp160 alone, signalled in CD4⁺ T cells. In addition, the sCD4/JR-FL gp160 complex induced a calcium flux in CD8⁺ T cells (data not shown). JR-FL gp160 induced a signal in HEK cells transfected with CD4 and CCR5, but not in HEK cells transfected with only CCR5 (S.V., P. Murphy, A.S.F. and D.W., unpublished observations). CD4 contributes to envelope-mediated signal transduction through CCR5 by inducing conformational changes in the envelope that enhance its ability to interact with CCR5. This is in agreement with previous reports that binding of envelope to CCR5 is enhanced by co-incubation with sCD4 (refs. 9, 10).

T-tropic gp160 derived from NL4-3 did not produce a signal in CD4⁺ T cells (Fig. 2d), whereas SDF-1, the natural ligand for the T-tropic HIV co-receptor CXCR4, was a potent stimulus (Fig. 2e). Additionally, NL4-3 gp160 blocked a subsequent response to JR-FL gp160, suggesting that NL4-3 gp160, which has a high affinity for CD4 (ref. 11), inhibited the initial binding between CD4 and JR-FL. Treatment of B10 cells with NL4-3 resulted in a calcium flux (Table

Figure 1 M-tropic gp160 induces CCR5-mediated calcium signalling in B10 cells. **a**, B10 cells were stimulated with buffer (HBSS with Ca²⁺ and Mg²⁺, 1% FCS, 10 mM HEPES) and JR-FL gp160 (0.6 nM) or MIP-1 β (100 nM). **b**, B10 cells were sequentially stimulated with buffer, JR-FL gp160 (1.8 nM), MIP-1 β (100 nM), and MIP-1 α (100 nM); the discontinuity is a result of a computer reset during data collection. **c**, B10 cells were sequentially stimulated with buffer, MIP-1 β (100 nM) and JR-FL gp160 (0.6 nM). Indo-1-loaded B10 cells in a Time Zero module-FACS Vantage flow cytometer were stimulated at the indicated time points (arrows) with the indicated ligand. Ratio of the fluorescence of bound to free Indo-1 versus time is shown. Data are representative of 5 different experiments.

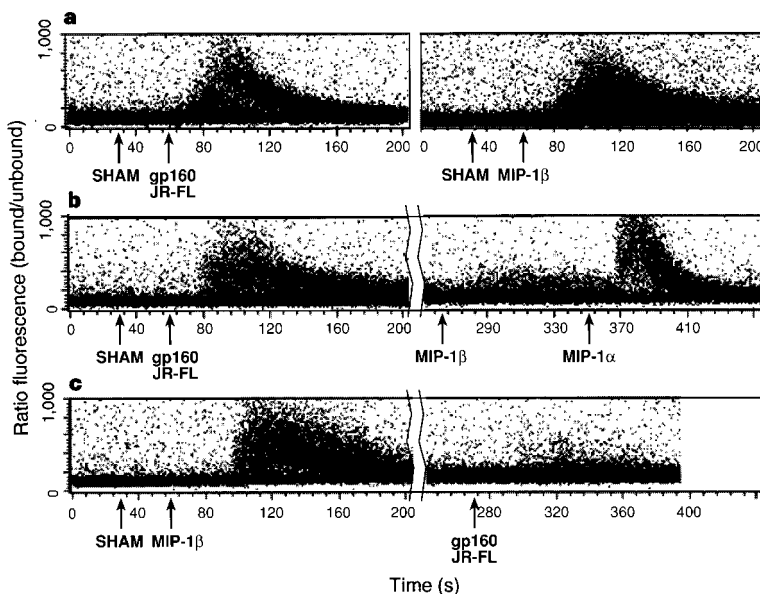


Table 1 Envelope proteins used in calcium flux assays

Ligand	Tropism	Induced signal in B10 cell line	Induced signal in CD4 ⁺ T cells (in PBL)
Control supt	—	—	—
gp160-JR-FL supt	M	+(0.3)*	+
gp160-NL4-3 supt	T	+(0.12)	—(6.0)†
Purified gp160-JR-FL	M	+(0.75)	+
gp160-PBj supt	M	+(0.04)	+
gp120-PBj	M	+(0.1)	+
gp120-Mac239	T	—	—(3.0)
gp120-IIIB	T	—	—(60)
gp160-Th014	NSI	+	+
gp160-CCa	SEQ	+	+
gp160-RW020-9 (clade A)	NSI	+	+

Supt, supernatant from envelope or control transfected CHO cells; M, M-tropic; T, T-cell-line tropic; NSI, non-syncytium inducing. The envelopes were shown to bind to CCR5 by displacement of MIP-1 β (ref. 16). SEQ, sequence of V3 suggests M-tropism.

* Concentration of protein (nM) that minimally stimulated B10 cells (>5% of cells responded). Similar concentrations also induced a signal in anti-CD3 plus IL-2 activated CD4⁺ T cells.

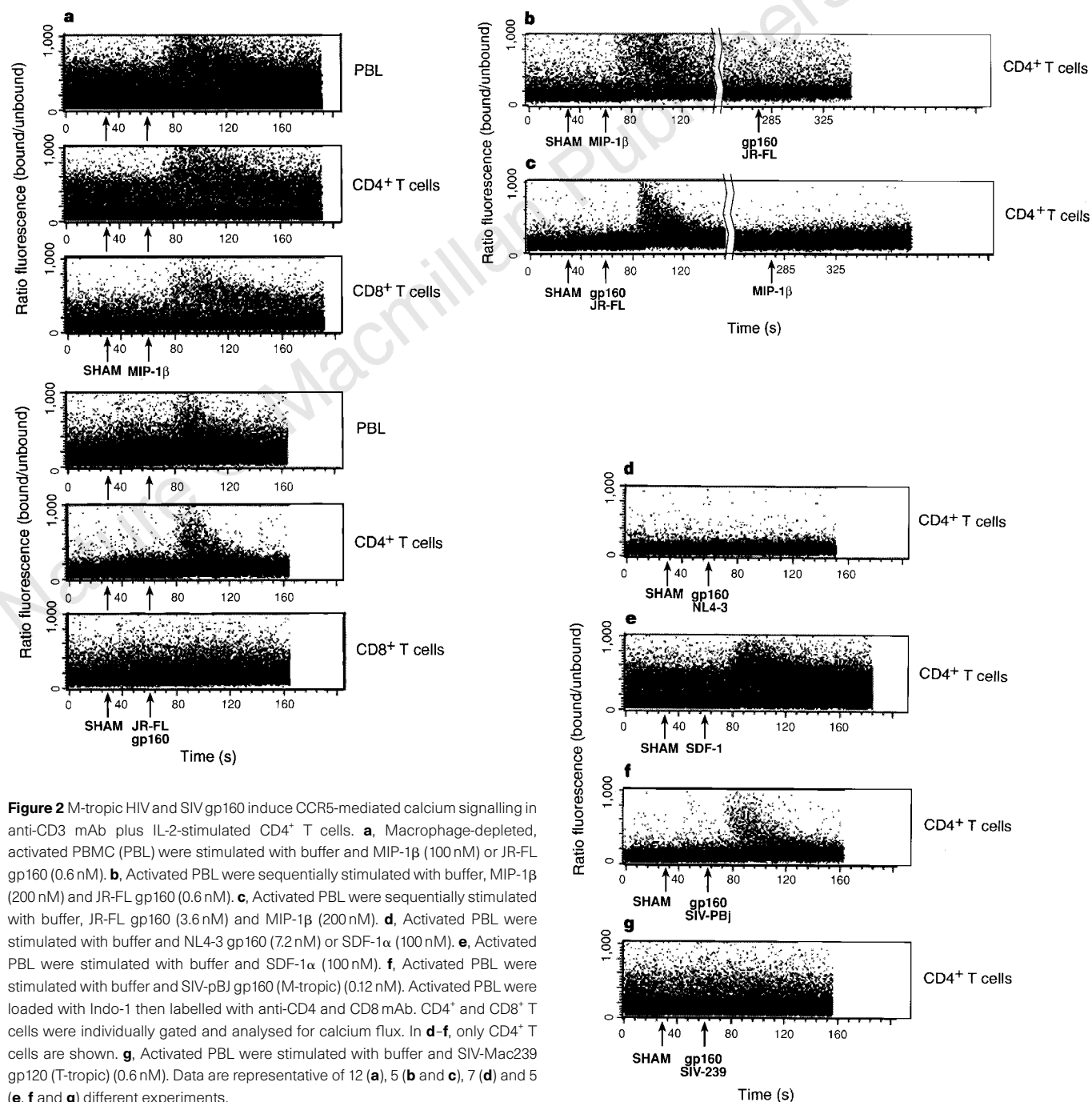
† Maximum concentration of protein (nM) used to determine a lack of signal induction.

Table 2 Partial inhibition of M-tropic gp160 induced calcium flux

Cells*	Stimulus	Percentage of cells responding†
B10	gp160-JR-FL supt (0.6 nM)	27.2
B10 + Leu3a	gp160-JR-FL (0.6 nM)	15.4
B10 + Leu3	gp160-JR-FL (0.6 nM)	25.7
CD4 ⁺ T cells	MIP-1 β (200 nM)	17.2
CD4 ⁺ T cells	gp160-JR-FL (1.2 nM)	7.8
CD4 ⁺ T cells plus Leu-3a	MIP-1 β (200 nM)	18.4
CD4 ⁺ T cells plus Leu-3a	gp160-JR-FL (1.2 nM)	3.1

* B10 and CD4⁺ T cells pretreated with Leu-3a, Leu-3 or medium were stimulated with the indicated ligand.

† The percentage of total B10 or CD4⁺ T cells responding to the stimulus was calculated using the Multitime software package.



1) that was sensitive to pertussis toxin; these cells did not respond to SDF-1 and expressed almost no CXCR4 messenger RNA, suggesting that gp160 derived from NL4-3 can signal through an alternative, non-CXCR4 receptor (data not shown). Thus, under certain conditions, as shown here with B10 cells, the NL4-3 envelope can induce a calcium flux. The failure of NL4-3 to induce calcium flux in CD4⁺ T cells may reflect a common distinction between M-tropic and T-tropic HIV envelopes. However, this recombinant envelope was derived from a highly adapted tissue-culture isolate of HIV-1 and may not represent naturally occurring T-tropic isolates.

Unlike HIV-1, both M-tropic and T-tropic SIVs use CCR5 as an entry cofactor^{12,13}. We used recombinant envelopes from the M-tropic SIV PBj1.9 and a recombinant envelope from the T-tropic SIV Mac239. These recombinant proteins bind directly to CCR5 on human CD4⁺ T cells, as measured by displacement of anti-CCR5

antibodies (J.A., A.R., A.S.F., unpublished observations) or MIP-1 β (ref. 10). Both the M-tropic (PBj) SIV gp160 (Fig. 2f) and gp120 (Table 1) stimulated CD4⁺ T cells to flux calcium, whereas the T-tropic SIV gp120 (Mac239) did not (Fig. 2g). Initial treatment with MIP-1 β desensitized cells to PBj gp160-induced calcium flux (data not shown). These results indicate that binding to CCR5 does not necessarily lead to calcium flux; rather, a degree of specificity in the interaction between envelope and CCR5 is required in order to observe signal transduction. These results, taken in the context of previous studies identifying the domains of CCR5 utilized by M-tropic and T-tropic SIVs (refs 12, 13), may help to clarify which domains of CCR5 are directly involved in and required for envelope-mediated signal transduction.

We investigated whether cells that signalled in response to gp160 would also chemotax towards HIV envelope protein. B10 cells chemotaxed to RANTES, MIP-1 α , MIP-1 β , and to M-tropic HIV gp160 (data not shown). Both the CD4⁺ and CD8⁺ cells in anti-CD3 plus IL-2-stimulated PBMC chemotaxed to MIP-1 β (Fig. 3a), RANTES, SDF-1 β and MIP-1 α (data not shown). In contrast, only CD4⁺ cells chemotaxed to M-tropic HIV JR-FL gp160 (Fig. 3b) and SIV PBj gp160 (Fig. 3c). No chemotaxis was observed in response to T-tropic HIV gp160 or SIV gp120 (Fig. 3d).

We have shown that five M-tropic envelopes (four HIV and one SIV) signal through CCR5 and induce chemotaxis at concentrations similar to those reported in the plasma of HIV-infected individuals¹⁴. Two T-tropic envelopes (one HIV and one SIV) did not induce a signal through their respective co-receptors. Both M-tropic HIV and SIV, which share relatively little sequence homology, use CCR5 for entry and are each capable of inducing a signal. In contrast, although both M-tropic and T-tropic SIV use CCR5 as an entry cofactor^{12,13,15}, only M-tropic SIV delivers a signal, as measured by calcium flux.

HIV disease is characterized by persistent immune activation². In this regard, envelope-mediated signalling through CCR5 may contribute directly or indirectly to this heightened state of activation. Such effects may be enhanced in the microenvironment of infected lymphoid tissue where target cells encounter, in addition to infectious virions, defective virions and soluble envelope protein. Because viral replication occurs most efficiently in activated cells², signals delivered through CCR5 have the potential to enhance the efficiency of HIV replication (ref. 16; A. Kinter, H. Moriuchi and A.S.F., unpublished observations). Finally, envelope-mediated cell signalling may contribute further to the pathogenesis of HIV disease by inducing the migration and accumulation of activated cells to sites of viral replication, thereby promoting the propagation and spread of virus.

Methods

Cells and reagents. B10 cells (from J. R. Yannelli) were periodically stimulated with phytohaemagglutinin (PHA; Sigma) and allogeneic, irradiated (3,000 rads) PBMC in RPMI 1640 (Biofluids) supplemented with 10% FCS (Sigma) and IL-2 (Chiron). PBMC were obtained by Ficoll-Hypaque centrifugation and stimulated with anti-CD3 (1:4,000 dilution of ascites) antibody OKT3 (ATCC) and IL-2 (20 U ml⁻¹). HIV-IIIB gp120 was from Advanced Biotechnologies. HIV-JR-FL, RW020-9, Th014 (ref. 10), CCA (ref. 17), NL4-3, SIV-PBj 1.9 and Mac239 (ref. 18) envelope proteins were obtained as described. Reagents used to generate these envelope proteins were obtained from the AIDS Reference Reagent program. HIV-blocking (Leu-3 α) and non-blocking (Leu-3) and anti-CD4 mAb were obtained from Becton-Dickinson; MIP-1 β , MIP-1 α , RANTES and SDF-1 α and β (R&D) were used at a final concentration of 100–200 nM.

Calcium signalling. The refined method of multiparameter FACS analysis of T-cell subsets and calcium flux will be described elsewhere (R.L.R., M. Park, F. Liao, R.S., D. Stephany and J.F.). Briefly, PBMC were loaded with indo-1 acetoxymethyl ester in the presence of pleuronic detergent (final concentrations, 10 μ M and 300 μ g ml⁻¹, respectively; Molecular Probes) at 30 °C for 45 min. To remove contaminating monocyte/macrophages, magnetic beads

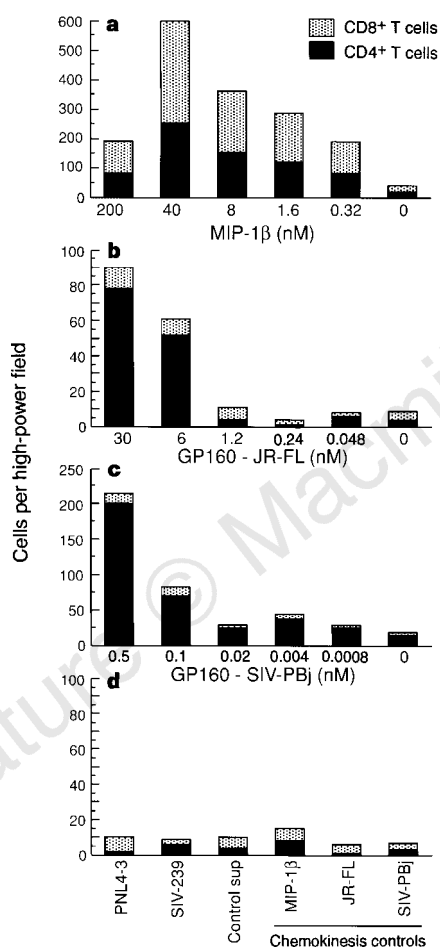


Figure 3 HIV and SIV M-tropic but not T-tropic envelopes are chemotactic for activated CD4⁺ T cells. Serial dilutions of envelopes derived from various HIV or SIV isolates or MIP-1 β were placed below 2- μ m filters. Labelled activated PBL were placed on top of the filters. Three hours later, the numbers of fluorescein (CD4⁺) and phycoerythrin (CD8⁺) labelled cells per high-power field in the lower chamber of microchemotaxis chambers were counted by fluorescence microscopy. **a**, MIP-1 β was chemotactic for both CD4⁺ and CD8⁺ T cells. Supernatants containing JR-FL, SIV-PBj and NL4-3 gp160 or SIV-Mac239 gp120 from transfected CHO cells were used in experiments shown in **b**, **c** and **d**, respectively. **b**, gp160 from JR-FL, and **c**, SIV-PBj were chemotactic for CD4⁺ T cells but not CD8⁺ T cells. **d**, gp160 derived from NL4-3 and gp120 derived from SIV Mac239 were not chemotactic for either CD4⁺ or CD8⁺ T cells. Control supernatant was obtained from mock-transfected CHO cells. **d**, Chemokinesis controls had MIP-1 β or gp160 added to both top and bottom wells. Each sample was run in triplicate and 5 high-power fields per sample were counted under fluorescence and light microscopy.

coated with anti-CD14 mAb (Dyna) were added to the cell suspension during Indo-1 loading. After washing, PBL were stained with fluorochrome-conjugated mAb (anti-CD4-Cy5PE (Sigma) and anti-CD8-PE (Pharmingen)) for 15 min at room temperature. All mAb were used at saturating concentrations.

Calcium flux was analysed with a Becton-Dickinson FACSVantage dual laser flow cytometer modified with a ratio offset (Becton-Dickinson) and a Time zero injection module (Cytex). Aliquots of PBL were warmed to 37°C for 3 min before analysis. The percentage of cells responding was determined using the software Multitime (Phoenix).

Chemotaxis. PBMC obtained 5–18 days after anti-CD3 stimulation were stained with CD8-PE (Pharmingen) and CD4-FITC (Biosource). 50,000 cells were placed above 2- or 3- μ m filters overlying medium, MIP-1 β or gp160 in serial dilutions in 96-well microchemotaxis chambers (Neuropore). Chemokinesis controls had MIP-1 β or gp160 in both the upper and lower wells. Cells were counted under light and fluorescence microscopy.

Received 9 May; accepted 17 September 1997.

- Murphy, P. M. Chemokine receptors: structure, function and role in microbial pathogenesis. *Cytokine Growth Factor Rev.* **7**, 47–64 (1996).
- Fauci, A. S. Host factors and the pathogenesis of HIV-induced disease. *Nature* **384**, 529–534 (1996).
- D'Souza, M. P. & Harden, V. A. Chemokines and HIV-1 second receptors. *Nature Med.* **2**, 1293–1300 (1997).
- Oravecz, T., Pall, M. & Norcross, M. A. Beta-chemokine inhibition of monocytotropic HIV-1 infection. Interference with a postbinding fusion step. *J. Immunol.* **157**, 1329–1332 (1996).
- Atchison, R. E. *et al.* Multiple extracellular elements of CCR5 and HIV-1 entry: dissociation from response to chemokines. *Science* **274**, 1924–1926 (1996).
- Liao, F. *et al.* STRL33, a novel chemokine receptor-like protein, functions as a fusion cofactor for both macrophage-tropic and T cell line-tropic HIV-1. *J. Exp. Med.* **185**, 2015–2023 (1997).
- Deng, H., Unutmaz, D., KewalRamani, V. N. & Littman, D. R. Expression cloning of new receptors used by simian and human immunodeficiency viruses. *Nature* **388**, 296–300 (1997).
- Alkhatib, G., Liao, F., Berger, E., Farber, J. M. & Peden, K. W. C. A new SIV co-receptor, STRL33. *Nature* **388**, 238 (1997).
- Wu, L. *et al.* D4-induced interaction of primary HIV-1 gp120 glycoproteins with the chemokine receptor CCR-5. *Nature* **384**, 179–183 (1996).
- Trikola, A. *et al.* CD4-dependent, antibody-sensitive interactions between HIV-1 and its co-receptor CCR-5. *Nature* **384**, 184–187 (1996).
- Kozak, S. L. *et al.* CD4, CXCR-4, and CCR-5 dependencies for infections by primary patient and laboratory-adapted isolates of human immunodeficiency virus type 1. *J. Virol.* **71**, 873–882 (1997).
- Edinger, A. L. *et al.* Differential utilization of CCR5 by macrophage and T cell tropic simian immunodeficiency virus strains. *Proc. Natl Acad. Sci. USA* **94**, 4005–4010 (1997).
- Chen, Z., Zhou, P. & Marx, P. A. Genetically divergent strains of simian immunodeficiency virus use CCR5 as a coreceptor for entry. *J. Virol.* **71**, 2705 (1997).
- Oh, S.-K. *et al.* Identification of HIV-1 envelope glycoproteins in the serum of AIDS and ARC patients. *J. AIDS* **5**, 251–256 (1992).
- Mori, K., Ringler, D. J. & Desrosiers, R. C. Restricted replication of simian immunodeficiency virus strain 239 in macrophages is determined by env but is not due to restricted entry. *J. Virol.* **67**, 2807–2814 (1993).
- Taub, D. D., Turcovski-Corralles, S. M., Key, M. L., Longo, D. L. & Murphy, W. J. Chemokines and T lymphocyte activation: I. Beta chemokines costimulate human T lymphocyte activation *in vitro*. *J. Immunol.* **156**, 2095–2103 (1996).
- McElrath, M. J. *et al.* Human immunodeficiency virus type 1 infection despite prior immunization with a recombinant envelope vaccine regimen. *Proc. Natl Acad. Sci. USA* **93**, 3972–3977 (1996).
- Mossman, S. P. *et al.* Protection against lethal simian immunodeficiency virus SIVsmmPBj14 disease by a recombinant Semliki Forest virus gp160 vaccine and by a gp120 subunit vaccine. *J. Virol.* **70**, 1953–1960 (1996).

Correspondence and requests for materials should be addressed to D.W. (e-mail: dweissman@nih.gov).

Polyisoprenyl phosphates in intracellular signalling

Bruce D. Levy*, Nicos A. Petasis† & Charles N. Serhan*

* Center for Experimental Therapeutics and Reperfusion Injury, Department of Anesthesia, Brigham and Women's Hospital and Harvard Medical School, 75 Francis Street, Boston, Massachusetts 02115, USA

† Department of Chemistry, University of Southern California, Los Angeles, California 90089, USA

In response to environmental stimuli, leukocyte membrane remodelling generates biologically active lipids that can serve as both intra- and extracellular mediators¹. There are several classes of lipids that can mediate inflammatory reactions.¹ We report here on a new intracellular lipid signal that regulates oxygen-radical formation in neutrophils, a key response in microbial killing, inflammation and tissue injury. Screening of neutrophil-derived extracts rich in phosphorylated, non-saponifiable lipids revealed a potent inhibitor of superoxide anion (O₂⁻) production. Structural

analysis of biologically active fractions gave four major phosphorylated lipids: most abundant was presqualene diphosphate (PSDP). Upon activation of neutrophil receptors, PSDP and its monophosphate form, presqualene monophosphate (PSMP), undergo rapid remodelling. At submicromolar concentrations, PSDP but not PSMP inhibit O₂⁻ production by human neutrophil cell-free oxidase preparations. We prepared PSDP and PSMP by total organic synthesis and matched both the physical properties and biological activity of the neutrophil-derived compounds. Our results indicate that PSDP, a recognized intermediate of cholesterol biosynthesis², is present in immune effector cells and is a potent regulator of the cellular response in host defence.

Receptor–ligand interactions result in the elaboration of both positive and negative signals, the timing and balance of which determines cellular responses¹. This phenomenon is exemplified by formyl-methionyl-leucyl-phenylalanine (FMLP) binding to its receptor on neutrophils, which causes the subsequent generation of intracellular signals that stimulate (for example, by intracellular calcium influx) as well as inhibit (for example, by formation of cyclic AMP) neutrophils³. Although several classes of intracellular lipids activate leukocytes during inflammation and reperfusion injury, few counterregulatory lipid-derived mediators have been identified⁴. To this end, we tested the capacity of neutrophil lipid extracts to regulate production of the superoxide anion. Neutrophil-derived lipid extracts triggered the generation of substantial amounts of superoxide anion (2.71 nmol O₂⁻ mg⁻¹ protein min⁻¹) by cell sonicates (Fig. 1a). We looked for regulatory signals by eliminating intact phospholipids with ester and/or amide bonds by saponification hydrolysis before extraction and chromatography on CC4 silica^{4,5}. The initial chloroform:methanol (2:1, v/v) elutions (fraction A) contained a mixture of fatty acids and neutral lipids that generated superoxide anion at similar rates (2.41 nmol mg⁻¹ protein min⁻¹). O₂⁻ production in these experiments was comparable to that found with known lipid activators in this system, namely arachidonic acid (75 μ M (C20:4), 1.38 nmol mg⁻¹ protein min⁻¹) and phosphatidic acid (100 μ M, 1,2-dicapryl-*sn*-glycero-3-phosphate (C10:0), 1.34 nmol mg⁻¹ protein min⁻¹), and was consistent with reported values^{6,7} (Fig. 1a). In contrast, subsequent elutions with chloroform:methanol:water (10:10:3, v/v) (fraction B), containing phosphorylated lipids that resist saponification, stimulated significantly less production of superoxide anion (0.62 nmol mg⁻¹ protein min⁻¹, *P* < 0.05) than fraction A, suggesting that fraction B carried an inhibitor. Furthermore, lipids present in fraction B potently inhibited O₂⁻ production stimulated by either phosphatidic acid (52%; Fig. 1b) or arachidonic acid (56%; see Supplementary information). These results indicate that material in fraction B counteracts intracellular lipid signals (that is, phosphatidic acid and arachidonic acid) operating in the initial activation of the neutrophil oxidase. Our findings are consistent with the biological properties of a proposed intracellular modulator of the oxidase which has remained elusive⁸.

To identify this inhibitor, fraction B was isolated from stimulated neutrophils and examined after thin-layer chromatography (TLC) (*n* = 3). Lipids from sequential 5-mm fractions (from origin to solvent front) were eluted and assessed for phosphorus content and activity. Regions at 10–15 mm from the origin carried phosphorus (11.3 nmol phosphorus per 10⁸ cells) and gave ~52% inhibition of phosphatidic acid-stimulated production of superoxide anion (Fig. 1c). To characterize fraction B's active material, lipids were chromatographed in parallel and stained with iodine to identify double bonds and molybdenum blue to visualize phosphorus: four principal compounds were seen (Fig. 1d). To analyse the structure, compound II was initially tested directly by gas chromatography/mass spectrometry (GC/MS), giving base peaks at *m/z* 69 with a fragmentation pattern of repeating C₅H₈ (*m/z* 68) units, ions that are characteristic of isoprenoids⁹ (see Supplementary information). *R_f* values in this TLC system were therefore determined for iso-

prenoids, including isopentenyl diphosphate, geranyl diphosphate, geranylgeranyl diphosphate (GGDP), dolichyl monophosphate (DOL-MP), isoprenyl alcohols, squalene or cholesterol (Fig. 1e). Compounds I and III (R_f values of 0.05 ± 0.01 and 0.17 ± 0.01) matched farnesyl diphosphate (FDP) and farnesyl monophosphate (FMP), respectively. Compounds II and IV carried the most phosphorus per cell and gave R_f values (0.10 ± 0.01 and 0.25 ± 0.02) that differed from these isoprenoids (Fig. 1d, e). Based on R_f values and the relationship between phosphorus content and mass (measured by scanning densitometry of charred TLC plates; see Supplementary information, Table 1), compound II was found to contain two phosphates and compounds III and IV one phosphate each. Together, these results indicate that four main phosphorylated lipids which resist saponification are present in neutrophils and have physical properties consistent with those of polyisoprenyl phosphates.

To identify the four compounds before assessing their biological activity, each was isolated and used for GC/MS, either by direct injection or after conversion to the Me_3Si derivatives. All four lipids gave the characteristic fragmentation pattern of isoprenoids⁹ (Supplementary information, Table 2). Compounds I and III gave retention times and prominent ions consistent with FDP and FMP which were confirmed by comparison with authentic material (Supplementary Information, Table 2). Compound II gave a fragmentation pattern resembling that of squalene (molecular ion (M^+), 410 a.m.u.), but gave ions of $m/z > 410$ and a different retention from squalene (Fig. 2a, b); combined with compound II's properties on TLC, this indicated a phosphorus-containing isoprenoid with $\geq 6 \text{ C}_5\text{H}_8$ units. To distinguish between a di- and monophosphorylated isoprenoid, compound II was treated with acid, which liberates 80% of the phosphate from polyisoprenyl diphosphates

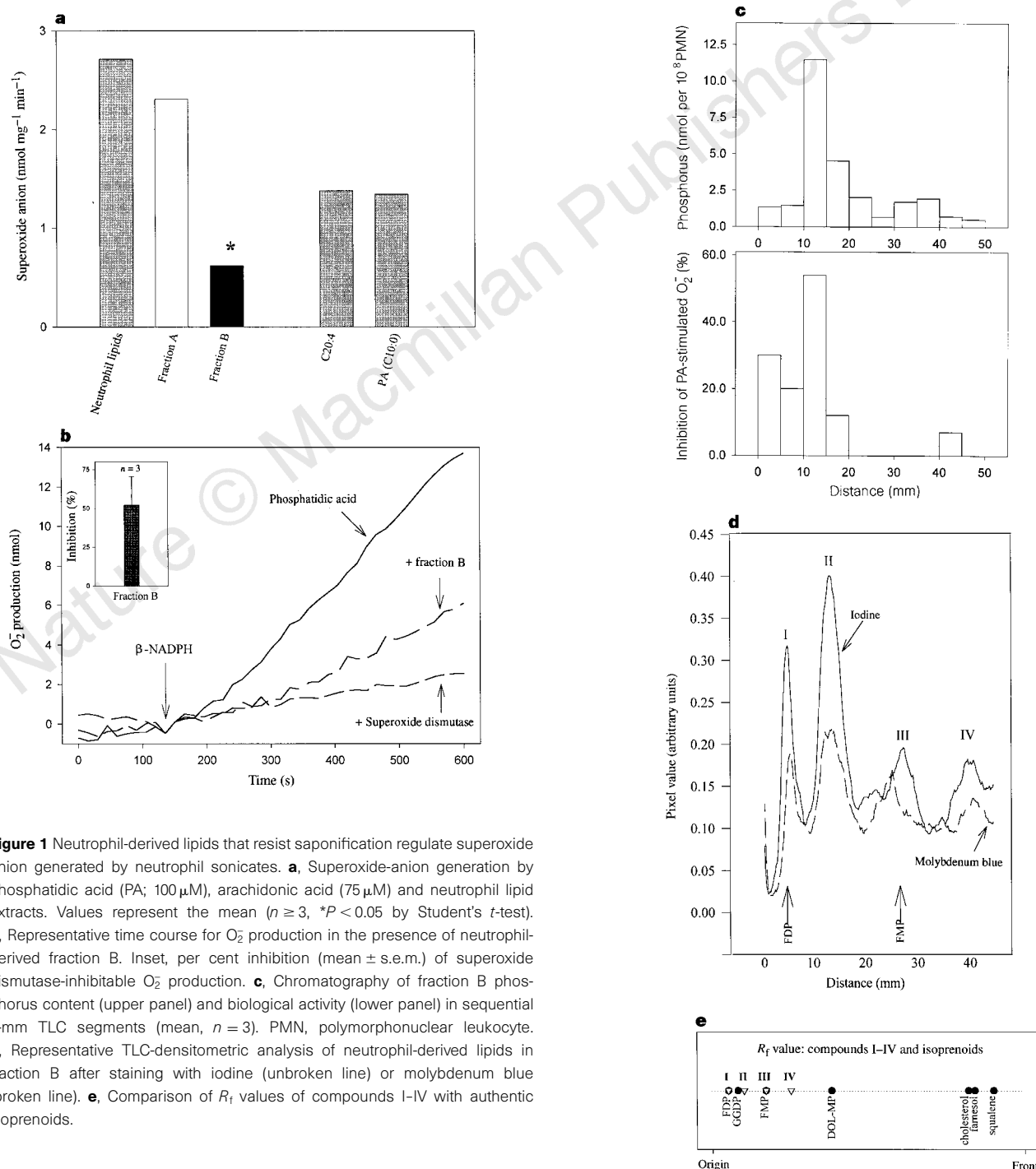


Figure 1 Neutrophil-derived lipids that resist saponification regulate superoxide anion generated by neutrophil sonicates. **a**, Superoxide-anion generation by phosphatidic acid (PA; 100 μM), arachidonic acid (75 μM) and neutrophil lipid extracts. Values represent the mean ($n \geq 3$, $*P < 0.05$ by Student's t -test). **b**, Representative time course for O_2^- production in the presence of neutrophil-derived fraction B. Inset, per cent inhibition (mean $\pm$ s.e.m.) of superoxide dismutase-inhibitable O_2^- production. **c**, Chromatography of fraction B phosphorus content (upper panel) and biological activity (lower panel) in sequential 5-mm TLC segments (mean, $n = 3$). PMN, polymorphonuclear leukocyte. **d**, Representative TLC-densitometric analysis of neutrophil-derived lipids in fraction B after staining with iodine (unbroken line) or molybdenum blue (broken line). **e**, Comparison of R_f values of compounds I-IV with authentic isoprenoids.

but not from polyisoprenyl monophosphates⁵; here, it converted compound II into squalene (Fig. 2b, c). Compound IV's retention and fragmentation pattern in GC/MS were similar to those of compound II, but compound IV was not hydrolysed by acid. No molecular ions were evident for either compound II or IV—consistent with the GC/MS analysis of polyisoprenyl phosphates^{10,11}. Based on these properties as well as evidence from material prepared by total organic synthesis (see below), we assigned compound II as presqualene diphosphate (PSDP) (Fig. 2a) and compound IV as presqualene monophosphate (PSMP).

In yeast, plants and mammalian liver, squalene synthase condenses two molecules of farnesyl diphosphate to form a cyclopropylcarbinyl intermediate termed presqualene diphosphate (PSDP) and, in the presence of NADPH, squalene.^{12–14} Human neutrophils generate squalene from endogenous sources, but lack the enzymes required for further conversion to cholesterol¹⁵, and neither PSDP nor PSMP have been identified in neutrophils. Compounds I, II, III and IV represented 0.1, 1.8, 0.4 and 0.5%, respectively, of the total phosphorylated lipid fraction (92.4 nmol phosphorus per 10⁷ unstimulated cells; *n* = 3). Thus, for a neutrophil of ~250 μm² in volume, the total concentration of compounds I–IV would be ~100 nM, not accounting for the compounds' intracellular distribution. Together, these results indicate that neutrophil lipids that resist saponification include a series of

biosynthetically related polyisoprenyl phosphates.

Upon receptor–ligand interaction, do compounds I–IV rapidly remodel in a way consistent with signalling events? Unstimulated neutrophils, exposed to [γ -³²P]-ATP to allow quantification and identify phosphate turnover, selectively incorporated radiolabel into compounds II (PSDP) and IV (PSMP) (Fig. 3). FMLP, the bacterial peptide analogue, stimulates superoxide anion generation in neutrophils and increases carboxyl-methylation of prenylated proteins through interaction with a G-protein-coupled, seven-transmembrane-spanning FMLP receptor^{16,17}. Exposure of neutrophils to FMLP resulted in changes in the phosphorus content of compounds II and IV which were rapid (within 60 s) and reciprocal (decrements in PSDP paralleled increments in PSMP) (Fig. 3, inset). Cytokines, such as granulocyte/monocyte-colony-stimulating factor (GM-CSF), prime neutrophils for inflammatory responses to FMLP by interacting with receptors whose signalling mechanisms are still not clear¹⁸. GM-CSF enhanced neutrophil ³²P incorporation into both compound II (45%) and compound IV (84%) (Supplementary information). Thus, remodelling PSDP and PSMP temporally overlaps the initial rate of O₂⁻ production triggered by receptor activation, and physiological stimuli increase PSDP and PSMP phosphate turnover in neutrophils.

To test the impact of neutrophil-derived compounds II and IV on the production of superoxide anions, isolated lipids were introduced by

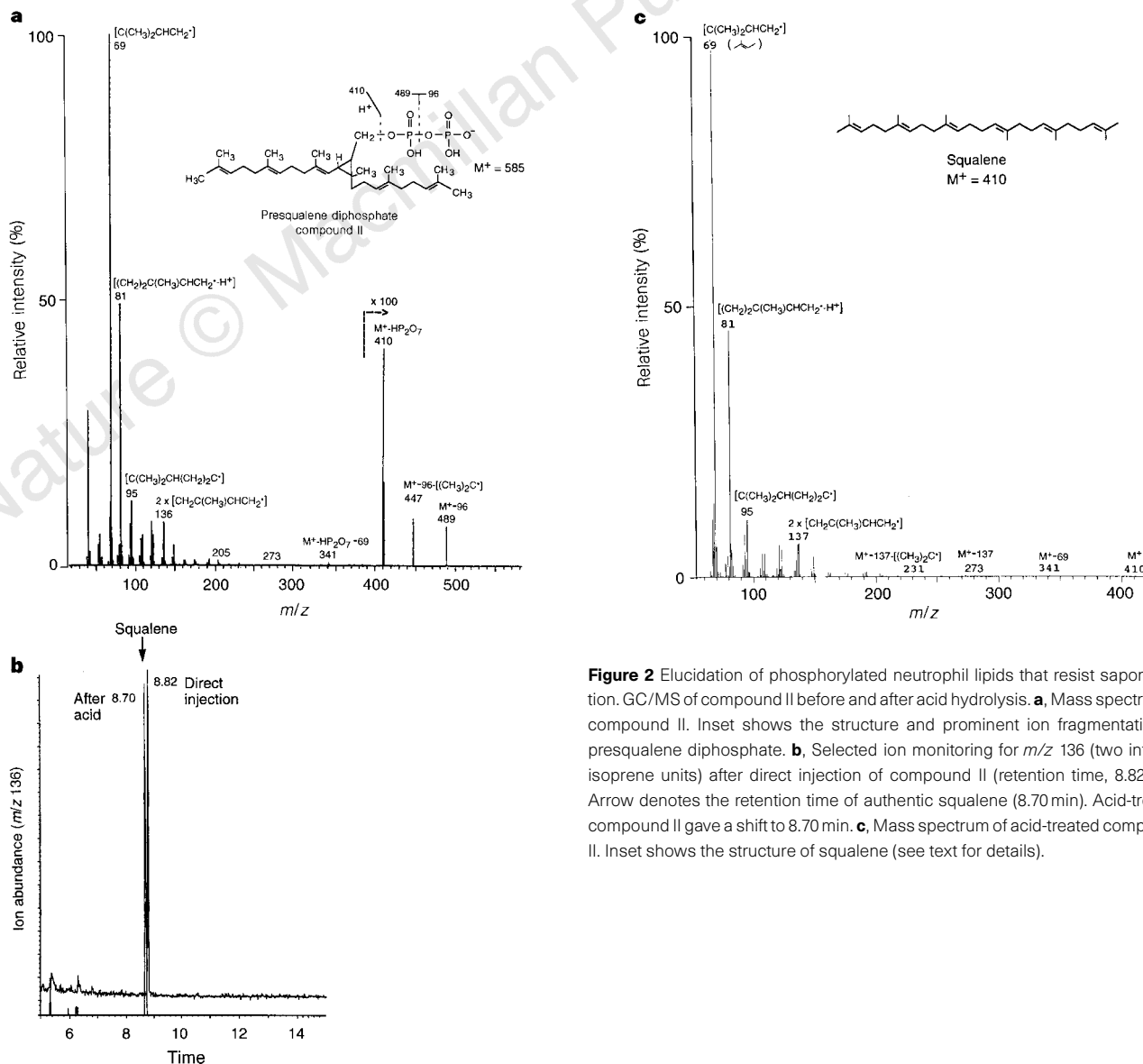


Figure 2 Elucidation of phosphorylated neutrophil lipids that resist saponification. GC/MS of compound II before and after acid hydrolysis. **a**, Mass spectrum of compound II. Inset shows the structure and prominent ion fragmentation of presqualene diphosphate. **b**, Selected ion monitoring for *m/z* 136 (two internal isoprene units) after direct injection of compound II (retention time, 8.82 min). Arrow denotes the retention time of authentic squalene (8.70 min). Acid-treated compound II gave a shift to 8.70 min. **c**, Mass spectrum of acid-treated compound II. Inset shows the structure of squalene (see text for details).

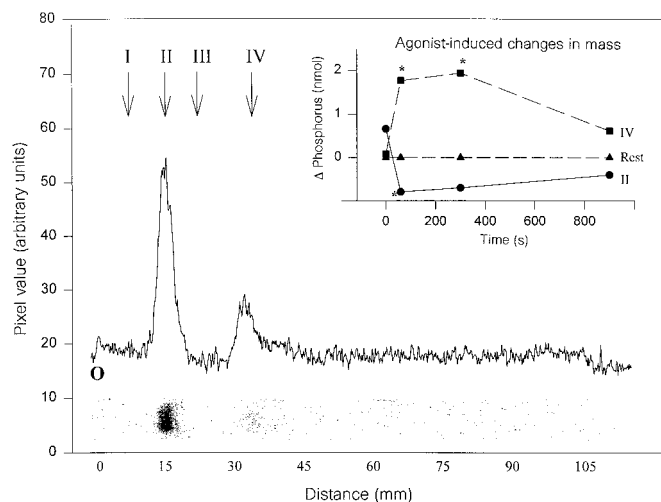


Figure 3 Receptor activation triggers remodelling of compound II and compound IV. Representative phosphorimager TLC profile of ^{32}P -content in neutrophil extracts from 20×10^6 cells per ml incubated with $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ ($40 \mu\text{Ci ml}^{-1}$) in the absence of an agonist (representative of $n = 6$, $d = 12$). Inset, time course with FMLP (10^{-7} M) stimulated (20×10^6 neutrophils per ml) changes in inorganic phosphorus content of compounds II (●) and IV (■) (without agonist ▲). Values are the mean of $n = 2$, $d = 4$. $*P < 0.03$ by Student's paired t -test. O, origin.

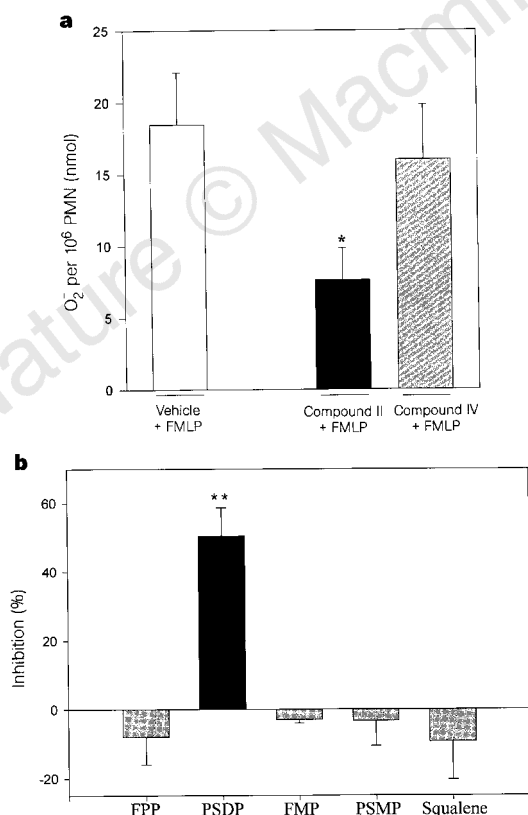
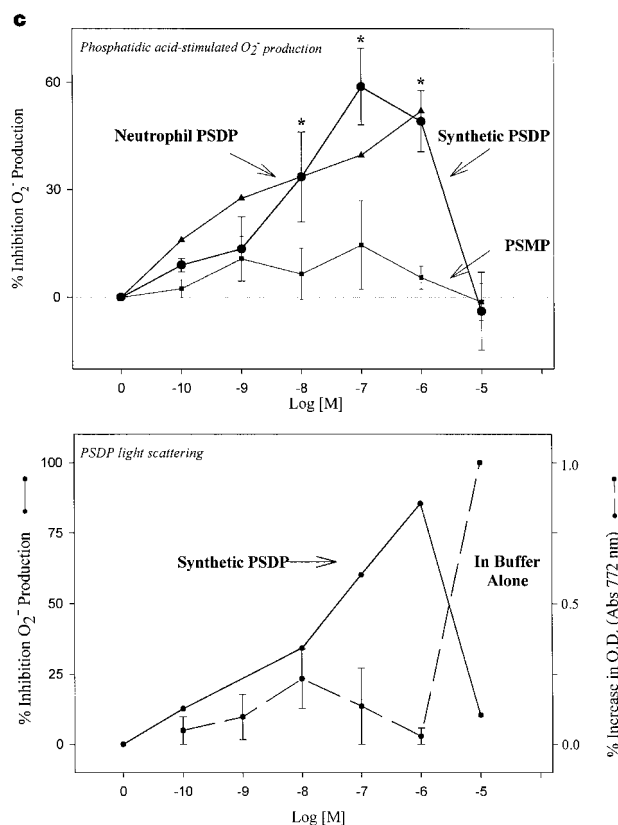


Figure 4 PSDP selectively inhibits neutrophil superoxide anion generation. Impact of neutrophil-derived compounds II (PSDP) and IV (PSMP) and closely related isoprenoids (10^{-6} M) on O_2^- stimulated by **a**, FMLP (10^{-7} M), and **b**, PMA (10^{-7} M) with electroporated neutrophils (PMN). **c**, Direct comparison of synthetic (●) and biologically derived PSDP (▲) and PSMP (■) in the concentration-dependent inhibition of phosphatidic acid-stimulated O_2^- production (upper panel). Phosphatidic acid ($100 \mu\text{M}$) gave $1.34 \pm 0.29 \text{ nmol } \text{O}_2^- \text{ mg}^{-1} \text{ min}^{-1}$. Per

electroporation before determining agonist-induced O_2^- production. PSDP (compound II) gave $\sim 60\%$ inhibition of FMLP-stimulated superoxide-anion generation, whereas O_2^- production in the presence of PSMP (compound IV) was unaltered (Fig. 4a). Neither compound alone stimulated O_2^- production (Supplementary information). As prenyl cysteine analogues disrupt G-protein interaction with activated FMLP receptors¹⁹, we tested the activity of PSDP in generating superoxide anion in neutrophil in the presence of phorbol-12 myristate 13-acetate (PMA), a potent agonist for O_2^- production that bypasses cell-surface receptors to stimulate protein kinase C and subsequent oxidase assembly¹⁷. PMA-activated superoxide-anion generation was inhibited by $\sim 45\%$ with PSDP ($1 \mu\text{M}$) in a concentration-dependent fashion, with significant inhibition at 10 nM (Fig. 4b, and Supplementary information). PMA-stimulated O_2^- production was not significantly blocked by PSMP, farnesyl diphosphate, or other closely related compounds of similar hydrophobicity (Fig. 4b). These results establish the structural basis for inhibition by PSDP because its monophosphate-containing form (PSMP) was devoid of activity at equimolar concentrations. In addition, our findings indicate that PSDP selectively inhibits O_2^- production at a site beyond the G-protein–receptor interaction. The inhibition of oxidase activity by PSDP, not seen with PSMP or other closely related isoprenoids, suggests a role for PSDP as a lipid-carrying signal, which contrasts with the structural roles of other isoprenoids, such as cholesterol, in the relay of signals as part of lipid rafts in cell membranes²⁰.

To confirm that PSDP carries the inhibitory activity of the biologically derived material, both PSDP and PSMP were prepared by total organic synthesis²¹. Each synthetic compound matched the



cent change in optical density for synthetic PSDP added to phosphate-buffered saline without cells ($n = 3$, mean $\pm$ s.e.m.) and comparison to inhibition of O_2^- production (lower panel). Absorbance (772 nm) with PSDP (10^{-5} M) was 7.84×10^{-4} . Values represent the mean $\pm$ s.e.m. for $n \geq 3$, $d \geq 6$. $**P < 0.01$, $*P < 0.05$ in Student's paired t -test compared to incubations in the presence of vehicle (0.2% ethanol) alone.

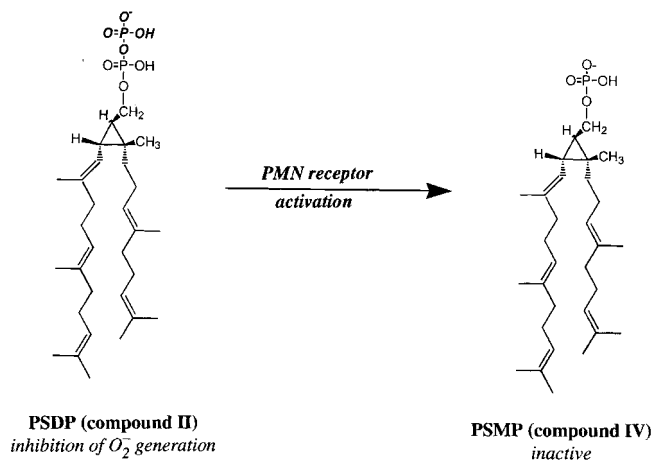


Figure 5 Scheme for polyisoprenyl phosphate remodelling.

physical and biological properties of neutrophil-derived PSDP (compound II) and PSMP (compound IV) (Fig. 4c; Supplementary information, Table 2). Again, synthetic PSDP, but not synthetic PSMP, inhibited superoxide anion generation in a concentration-dependent fashion in the presence of phosphatidic acid, a physiologically relevant endogenous stimulus (Fig. 4c). Inhibition ($\sim 60\%$) was maximal at 10^{-7} M PSDP. Submicromolar concentrations of PSDP significantly inhibited both FMLP (10^{-7} M)- and PMA (10^{-8} M)-stimulated oxidase activity in electroporated neutrophils (see Supplementary information). Concentrations of PSDP $> 10^{-6}$ M did not further inhibit O_2^- production (Fig. 4c). At 10^{-5} M, PSDP gave a bell-shaped dose-response curve, like that obtained with some other active lipids.¹ In this case, PSDP aggregates in buffer alone at 10^{-5} M, as evidenced by the increase in absorbance (Fig. 4c). Thus, above physiological concentrations, aggregation prevents PSDP from inhibiting O_2^- production. These findings indicate that neutrophil compounds II (PSDP) and IV (PSMP) share the physical and biological properties of synthetic PSDP and PSMP, and that nanomolar to micromolar PSDP (levels achieved during neutrophil activation and remodelling) regulates agonist-induced neutrophil responses during inflammation and host defence.

Isoprenoids are important intermediates in the generation of several classes of compounds, including sterols, retinoids, dolichols, ubiquinone and prenylated proteins^{4,12,22}. We have identified the principal polyisoprenyl phosphates in neutrophils, and show that, in response to receptor activation, two of these compounds, presqualene diphosphate and monophosphate, rapidly remodel. Also, physiological concentrations of presqualene diphosphate regulated cellular activities relevant during inflammation and tissue injury (Fig. 5). The finding that PSDP is a negative regulator in neutrophils does not prevent PSDP from stimulating actions with other cell types, as exemplified by the lipoxins, which inhibit neutrophil yet stimulate monocyte motility²³. Together, our results show that PSDP is a potent lipid-derived signal and provide evidence for isoprenoid remodelling and a link between cholesterol metabolism and immune function. □

Methods

Isolation of neutrophil lipids. Venous blood (~ 180 ml) was drawn by venipuncture from healthy volunteers who had not taken any medication for at least two weeks. Neutrophils were isolated from whole blood²³. Neutrophil lipids were saponified (10% KOH in methanol at 37°C for 30 min), extracted⁵ and separated by CC4 silica (~ 1 g per 10^8 cells) chromatography with sequential elutions of chloroform:methanol (2:1, v/v) (fraction A) and chloroform:methanol:water (10:10:3, v/v) (fraction B). Eluents were roto-evaporated *in vacuo*, concentrated in chloroform:methanol (2:1, v/v) and spotted on TLC plates for development (40–50 min, 25°C) in an environment saturated with chloroform:methanol:water (65:25:4, v/v).

For phosphorus determination and assay of activity in fraction B after TLC, sequential 5-mm segments of each lane were scraped and eluted with 1 ml of chloroform:methanol (2:1, v/v) four times. Samples were concentrated in a stream of N_2 and inorganic phosphorus was quantified²⁴. For evaluation of biological activity, aliquots of material were suspended in ethanol and vortexed immediately before assay.

To stain compounds for densitometry, TLC plates with material run in parallel were either exposed to sublimed iodine (1–5 min, 25°C), sprayed with 10% CuSO_4 in 8% phosphoric acid, and charred (100°C for 20 min, followed by 120°C for 10 min)²⁵, or sprayed directly with molybdenum blue:4.2 M sulphuric acid (1:1, v/v). R_f values were determined from the leading edge of each compound. On occasion, compounds II, III and IV appeared as doublets, suggesting the presence of closely related compounds such as isomers within each major peak.

After quantification by inorganic phosphorus determination²⁴, ~ 500 ng of phosphorus per each compound in hexane (2 μl) was taken to GC/MS (Hewlett-Packard, model 5890 GC and 5871A MS). To enhance detection, Me_3Si derivatives were produced by treatment with BSTFA, or compounds were treated with acid (0.01 M HCl in 50% methanol, pH 2.0 at 100°C for 20 min) to liberate pyrophosphate when present⁵. Diagnostic ions in Fig. 2a, namely m/z 489 ($\text{M}-96$), 447 $\text{M}-[96\cdot\text{C}(\text{CH}_3)_2]$, 410 $\text{M}-[(\text{HP}_2\text{O}_7)-(\text{C}(\text{CH}_3)_2\text{CHCH}_2)]$, 273 $\text{M}-[(\text{HP}_2\text{O}_7)-69\cdot(\text{CH}_2\text{C}(\text{CH}_3)\text{CHCH}_2)]$, 205 $\text{M}-[(\text{HP}_2\text{O}_7)-69\cdot(\text{CH}_2\text{C}(\text{CH}_3)\text{CHCH}_2 \times 2)]$, 136 $[(\text{CH}_2\text{C}(\text{CH}_3)\text{CHCH}_2 \times 2)]$, 95 $[(\text{CH}_3)_2\text{CH}(\text{CH}_2)_2\text{C}]$, 81 $[(\text{CH}_2\text{C}(\text{CH}_3)\text{CH}(\text{CH}_2)_2\text{H}^+)]$ and 69 [base peak; $(\text{CH}_3)_2\text{CCHCH}_2]$, were consistent with presqualene diphosphate (or, in Fig. 2c, squalene^{10,11}). (High-resolution mass-spectral data from synthetic PSDP established the molecular ion as 585 a.m.u.; see Supplementary information.)

Labelling of neutrophil PSDP and PSMP. Cells (20×10^6 per ml in phosphate-free HHBSS) were labelled with $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ (40 $\mu\text{Ci ml}^{-1}$, 90 min, 37°C), and saponified (10% KOH in methanol for 30 min at 37°C). Non-saponifiable lipids were extracted and separated by one-dimensional TLC (chloroform:methanol:water (65:25:4 (v/v) for 50 min at 25°C). Phosphoimaging (model 425E and integration software; Molecular Dynamics) was used to determine ^{32}P . As destruction of the TLC plate was not required for ^{32}P analysis, the same individual lanes were later exposed to iodine vapour and charred before densitometry. To determine FMLP (10^{-7} M)-induced changes in compound II (PSDP) and IV (PSMP) mass, radioactive compounds were eluted and inorganic phosphorus was measured for each compound.

Superoxide-anion generation. Neutrophil sonicates: To determine O_2^- production, post-nuclear supernatants from neutrophils after sonication (0.1–0.5 mg protein ml^{-1}) (ref. 6), 100 μM phosphatidic acid (1,2-dicapryl-*sn*-glycero-3-phosphate; Avanti Polar Lipids), 75 μM arachidonic acid and/or lipid extracts (0.1% ethanol) were added to cytochrome *c* (60 μM), sucrose (170 mM), EGTA (1 mM), FAD (10 μM), $\text{Na}_2\text{S}_2\text{O}_8$ (2 mM) and either superoxide dismutase (30 $\mu\text{g ml}^{-1}$) or H_2O at 0°C (as in ref. 26). After 2 min (37°C , pH 6.8) in a thermal-jacketed cuvette, $\beta\text{-NADPH}$ (200 μM) was added to initiate reactions and the absorbance at 550 nm was monitored at 15-s intervals or after 10 min ($n \geq 3$). Electroporated neutrophils: 4×10^6 neutrophils (5×10^6 cells per ml) were electroporated (1.65 kV cm^{-1} , 250 μF , infinite resistance; Invitrogen Electroporator II) at 0°C in the presence of 10^{-9} – 10^{-6} M compound II, IV, FDP, FMP, squalene, synthetic compound or vehicle (0.2% ethanol). After electroporation, 0.5×10^6 cells in cytochrome *c* (0.7 mg ml^{-1}) were exposed (for 10 min at 37°C and pH 7.45) to either cytochalasin *b* (5 $\mu\text{g ml}^{-1}$) and FMLP (10^{-7} M) or PMA (10^{-8} – 10^{-7} M) and the absorbance of the supernatants was monitored at 550 nm. Electroporation gave $\sim 50\%$ ($n = 10$) reduction in FMLP-stimulated NADPH oxidase compared to non-electroporated cells.

Preparation of synthetic materials. Synthetic PSDP and PSMP were prepared by total organic synthesis²¹ and their structures were confirmed by NMR (^1H , ^{13}C , ^{31}P) and high-resolution MS (see Supplementary information). Before assay, the *tert*-butyl ammonium salt of PSDP was converted to its anionic form by pH titration to 9.0 and elution from anion-exchange resin (DEAE-Sephacel) (2 ml resin, C:M:W 10:10:3, v/v; pH 9.0 for the mobile phase at 25°C) followed by TLC.

Received 23 June; accepted 8 September 1997.

- Serhan, C. N., Haeggström, J. Z. & Leslie, C. C. Lipid mediator networks in cell signaling: Update and impact of cytokines. *FASEB J.* **10**, 1147–1158 (1996).
- Jarstfer, M. B., Blagg, B. S. J., Rogers, D. H. & Poulter, C. D. Biosynthesis of squalene. Evidence for a

- tertiary cyclopropylcarbonyl cationic intermediate in the rearrangement of presqualene diphosphate to squalene. *J. Am. Chem. Soc.* **118**, 13089–13090 (1996).
3. Weissmann, G., Smolen, J. E. & Korchak, H. M. Release of inflammatory mediators from stimulated neutrophils. *N. Engl. J. Med.* **303**, 27–34 (1980).
 4. Adair, W. L. & Keller, R. K. Isolation and assay of dolichol and dolichyl phosphate. *Meth. Enzymol.* **111**, 201–215 (1985).
 5. Van Dessel, G. A. E., Lagrou, A. R., Hilderson, H. J. & Dierick, W. S. H. in *CRC Handbook of Chromatography* (eds Mukherjee, K. D., Weber, N. & Sherma, J.) 321–337 (CRC Press, Boca Raton, 1993).
 6. McPhail, L. C., Shirley, P. S., Clayton, C. C. & Synderman, R. Activation of the respiratory burst enzyme from human neutrophils in a cell-free system. Evidence for a soluble cofactor. *J. Clin. Invest.* **75**, 1735–1739 (1985).
 7. Agwu, D. E., McPhail, L. C., Sozzani, S., Bass, D. A. & McCall, C. E. Phosphatidic acid as a second messenger in human polymorphonuclear leukocytes. Effects on activation of NADPH oxidase. *J. Clin. Invest.* **88**, 531–539 (1991).
 8. Bokoch, G. M. & Knaus, U. G. Ras-related GTP-binding proteins and leukocyte signal transduction. *Curr. Opin. Hematol.* **1**, 53–60 (1994).
 9. Farnsworth, C. C., Gelb, M. H. & Glomset, J. A. Identification of geranylgeranyl-modified proteins in HeLa cells. *Science* **247**, 320–322 (1990).
 10. Popják, G., Edmond, J., Clifford, K. & Williams, V. Biosynthesis and structure of a new intermediate between farnesyl pyrophosphate and squalene. *J. Biol. Chem.* **244**, 1897–1918 (1969).
 11. Epstein, W. W. & Rilling, H. C. Studies on the mechanism of squalene biosynthesis. The structure of presqualene pyrophosphate. *J. Biol. Chem.* **245**, 4597–4605 (1970).
 12. Goldstein, J. L. & Brown, M. S. Regulation of the mevalonate pathway. *Nature* **343**, 425–430 (1990).
 13. Corey, E. J. & Volante, R. P. Application of unreactive analogs of terpenoid pyrophosphates to studies of multistep biosynthesis. Demonstration that “presqualene pyrophosphate” is an essential intermediate on the path to squalene. *J. Am. Chem. Soc.* **98**, 1291–1293 (1976).
 14. Mookhtiar, K. A., Kalinowski, S. S., Khang, D. & Poulter, C. D. Yeast squalene synthase. A mechanism for addition of substrates and activation by NADPH. *J. Biol. Chem.* **269**, 11201–11207 (1994).
 15. Shechter, I., Fogelman, A. M. & Popjak, G. A deficiency of mixed function oxidase activities in the cholesterol biosynthetic pathway of human granulocytes. *J. Lipid Res.* **21**, 277–283 (1980).
 16. Phillips, M. R. et al. Carboxyl methylation of ras-related proteins during signal transduction in neutrophils. *Science* **259**, 977–980 (1992).
 17. Baggiolini, M., Boulay, F., Badwey, J. A. & Curnutte, J. T. Activation of neutrophil leukocytes: chemotactant receptors and respiratory burst. *FASEB J.* **7**, 1004–1010 (1993).
 18. Gomez-Cambronero, J. & Sha’afi, R. I. in *Cell-Cell Interactions in the Release of Inflammatory Mediators* (eds Wong, P. Y.-K. & Serhan, C. N.) 35–71 (Plenum, New York, 1991).
 19. Scheer, A. & Gierschik, P. S-prenylated cysteine analogues inhibit receptor-mediated G protein activation in native human granulocyte and reconstituted bovine retinal rod outer segment membranes. *Biochemistry* **34**, 4952–4961 (1995).
 20. Simons, K. & Ikonen, E. Functional rafts in cell membranes. *Nature* **387**, 569–572 (1997).
 21. Rogers, D. H., Yi, E. C. & Poulter, C. D. Enantioselective synthesis of (+)-presqualene diphosphate. *J. Org. Chem.* **60**, 941–945 (1995).
 22. Quinn, M. T. Low-molecular-weight GTP-binding proteins and leukocyte signal transduction. *J. Leuk. Biol.* **58**, 263–276 (1995).
 23. Maddox, J. F. & Serhan, C. N. Lipoxin A₄ and B₄ are potent stimuli for human monocyte migration and adhesion: selective inactivation by dehydrogenation and reduction. *J. Exp. Med.* **183**, 137–146 (1996).
 24. Chen, P. S., Toribara, T. Y. & Warner, H. Microdetermination of phosphorus. *Analyt. Chem.* **28**, 1756–1758 (1956).
 25. Tou, J. & Dola, T. Leukotriene B₄ stimulation of an early elevation of phosphatidic acid mass in human neutrophils. *Lipids* **30**, 373–381 (1995).
 26. Bromberg, Y. & Pick, E. Unsaturated fatty acids stimulate NADPH-dependent superoxide production by cell-free system derived from macrophages. *Cell. Immunol.* **88**, 213–221 (1984).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from Mary Sheehan at the London editorial office of Nature.

Acknowledgements. This work was supported in part by the NIH (GM & NIDDK) and a basic discovery research grant from ONO Pharmaceuticals to C.N.S. and the NIH (GM) to N.A.P. B.D.L. was supported in part by a Paul Dudley White postdoctoral fellowship of the Massachusetts Affiliate, American Heart Association. We thank Valery V. Fokin for help with the synthesis and the NMR analyses of PSDP and PSMP.

Correspondence and requests for materials should be addressed to C.N.S. (e-mail address: CNSerhan@zeus.BWH.harvard.edu).

Calcium sensitization of smooth muscle mediated by a Rho-associated protein kinase in hypertension

Masayoshi Uehata, Toshimasa Ishizaki*, Hiroyuki Satoh, Takashi Ono, Toshio Kawahara, Tamami Morishita, Hiroki Tamakawa, Keiji Yamagami, Jun Inui, Midori Maekawa* & Shuh Narumiya*

Discovery Research (Tokyo), Yoshitomi Pharmaceutical Industries Ltd, Iruma, Saitama 358, Japan

* Department of Pharmacology, Faculty of Medicine, Kyoto University, Yoshida, Kyoto 606-01, Japan

Abnormal smooth-muscle contractility may be a major cause of disease states such as hypertension, and a smooth-muscle relaxant that modulates this process would be useful therapeutically. Smooth-muscle contraction is regulated by the cytosolic Ca²⁺ concentration and by the Ca²⁺ sensitivity of myofilaments¹: the

former activates myosin light-chain kinase and the latter is achieved partly by inhibition of myosin phosphatase^{1–3}. The small GTPase Rho and its target, Rho-associated kinase, participate in this latter mechanism *in vitro*^{4–6}, but their participation has not been demonstrated in intact muscles. Here we show that a pyridine derivative, Y-27632, selectively inhibits smooth-muscle contraction by inhibiting Ca²⁺ sensitization. We identified the Y-27632 target as a Rho-associated protein kinase, p160ROCK⁷. Y-27632 consistently suppresses Rho-induced, p160ROCK-mediated formation of stress fibres in cultured cells and dramatically corrects hypertension in several hypertensive rat models. Our findings indicate that p160ROCK-mediated Ca²⁺ sensitization is involved in the pathophysiology of hypertension and suggest that compounds that inhibit this process might be useful therapeutically.

We have developed a series of compounds that are potent relaxants of vascular and bronchial smooth muscles: Y-27632 is a representative compound (Fig. 1a) which potently inhibits phenylephrine-induced contraction of rabbit aortic strips, with a half-maximal inhibitory concentration (IC₅₀) of 0.7 μM (Fig. 1b); however, it had little effect on contraction induced by potassium chloride (IC₅₀ > 30 μM). Y-27632 also inhibited the contraction of pig coronary artery strips and guinea-pig trachea induced by various agonists such as phenylephrine, histamine, acetylcholine, serotonin, endothelin and a thromboxane agonist, U-46619 (IC₅₀ values of 0.3–1 μM; data not shown). These results indicate that Y-27632 is a specific and potent inhibitor of agonist-induced smooth-muscle contraction.

The effects of this compound were next examined on permeabilized strips of rabbit mesenteric artery. In these preparations, Ca²⁺-dependent contraction can be induced by an increase in the extracellular Ca²⁺ concentrations, whereas addition of GTPγS can evoke contraction through a Ca²⁺-sensitization mechanism⁸. Y-27632 inhibited GTPγS-induced contraction in a concentration-dependent way, like the inhibition of phenylephrine-induced contraction in intact strips, but it had no effect on Ca²⁺-induced contraction (Fig. 1c). Contraction in this system can also be evoked by inhibition of myosin phosphatase with a phosphatase inhibitor, calyculin A⁹. Y-27632 had no effect on this contraction up to 100 μM, whereas a myosin light-chain kinase inhibitor, ML-9 (ref. 10), abolished this contraction (Fig. 1c, bottom trace). These results indicate that Y-27632 inhibits smooth-muscle contraction selectively by inhibiting the Ca²⁺-sensitization mechanism.

To identify the molecular target of Y-27632, we did a radioligand binding assay using [³H]Y-30141, a ³H-labelled analogue of Y-27632, and found specific binding in the membrane as well as in the cytosol (Fig. 2a). Scatchard analysis in the membrane yielded a single class of binding site (K_d of 8 nM, B_{max} of 1.0 pmol mg^{−1} protein). This binding was efficiently displaced by a series of Y-27632 analogues as well as by various protein kinase inhibitors. Displacement potency correlated well with ability to inhibit phenylephrine-induced smooth-muscle contraction: a strong positive correlation was seen, with a slope close to unity (Fig. 2b). This binding protein is likely to be the target for Y-27632-induced inhibition of smooth-muscle contraction.

To isolate this binding protein, we used ¹²⁵I-labelled Y-35526, another Y-27632 analogue, as a photoaffinity ligand. Photoaffinity labelling with this compound gave several bands corresponding to relative molecular masses (M_r) of 150K–160K on SDS-polyacrylamide gel electrophoresis. This labelling was dependent on irradiation time and was inhibited by excess Y-30141 (Fig. 2c). A major 150K protein was purified and subjected to cyanogen bromide fragmentation. Partial amino-acid sequences of the 6.5K and 18K fragments were identified as XRKDFKIDRK (where X is N, D, A or G) and XLDIDKLFHV (X is A or V), respectively. These are identical to amino acids 1,011–1,020 and 1,167–1,176, respectively, of p160ROCK, a Rho-associated coiled-coil forming kinase⁷. Consistently, this 150K

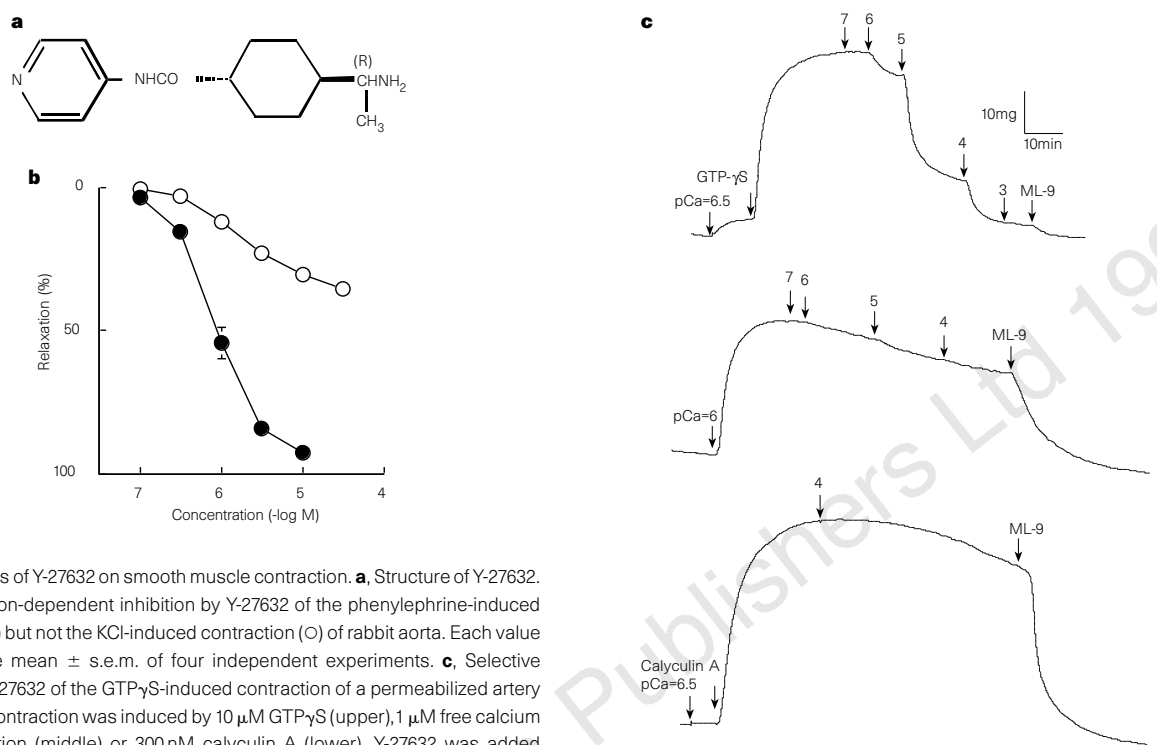


Figure 1 Effects of Y-27632 on smooth muscle contraction. **a**, Structure of Y-27632. **b**, Concentration-dependent inhibition by Y-27632 of the phenylephrine-induced contraction (●) but not the KCl-induced contraction (○) of rabbit aorta. Each value represents the mean $\pm$ s.e.m. of four independent experiments. **c**, Selective inhibition by Y-27632 of the GTP γ S-induced contraction of a permeabilized artery preparation. Contraction was induced by 10 μ M GTP γ S (upper), 1 μ M free calcium ion concentration (middle) or 300 nM calyculin A (lower). Y-27632 was added cumulatively at the concentrations shown by -log [molar concentration] above each trace. ML-9 was added at 300 μ M at the indicated times in each experiment.

protein reacted with the anti-p160ROCK antibody on western blot analysis (data not shown). To confirm this result, we tested the effects of Y-27632 and related compounds on the kinase activity of p160ROCK purified from human platelets⁷. All these compounds inhibited the kinase activity in a dose-dependent manner (Fig. 2d) and their inhibitory potency was well correlated with their ability to inhibit [³H]Y-30141 binding and agonist-induced contraction (data not shown). p160ROCK has an isoenzyme called either ROK α ¹¹, Rho kinase¹² or ROCK-II¹³, which is ~90% identical to p160ROCK in its kinase domain. We therefore expressed these kinases in COS cells and examined the inhibition by Y-27632 and related compounds on their activity. Identical profiles and sensitivities were found for the two kinases towards Y-27632, Y-30141, Y-30964 and HA1077. The K_i values of Y-27632 and related compounds for several kinases are compared in Table 1. Of the five compounds tested, Y-27632 showed the best selectivity: its affinity for p160ROCK was ~200 times or over 2,000 times higher than that for conventional protein kinase C (PKC) from rat brain or cAMP-dependent protein kinase, or for myosin light-chain kinase, respectively. The affinity for PKC ϵ , a protein kinase C isoenzyme, was a little higher but still 10–50-fold lower than that for p160ROCK (our unpublished observations). Y-27632 did not appreciably inhibit at 100 μ M concentration PAK¹⁴, a protein kinase activated by another Rho-family GTPase, Cdc42.

Rho works as a molecular switch for the induction of focal adhesions and stress fibres in cultured cells¹⁵, and p160ROCK¹⁶ or ROK α /Rho-kinase/ROCK-II^{17,18} work as Rho effectors in this process. We therefore tested whether Y-27632 could inhibit Rho-induced formation of these structures in cultured cells. Treatment of cells with Y-27632 abolished focal adhesions and stress fibres induced not only by overexpression of p160ROCK but also by activated Rho in 30 min after the addition (Fig. 3a, b). It also inhibited induction of stress fibres by lysophosphatidic acid in cultured fibroblasts (data not shown). Y-30141 inhibited induction of this phenotype with a 10-fold higher potency. In addition to stress fibre formation, Rho exerts a variety of actions in the cell, including transcriptional regulation by serum-response factor¹⁹; Y-27632, however, does not inhibit this transcriptional process (E. Sahai and R. Treisman, personal communication). Y-27632 did not inhibit Rac-induced membrane ruffling²⁰ in cultured cells (Fig. 3c). These findings show that Y-27632 can work as a specific inhibitor of p160ROCK in intact cells.

Our results indicate that Y-27632 acts on p160ROCK in intact smooth muscle to inhibit agonist-induced contraction. We next analysed the importance of this mechanism and its pathophysiological role in the regulation of smooth muscle tone *in vivo*. Y-27632 was administered *per os* to various hypertensive rat models. As shown in Fig. 4a, this compound significantly decreased the blood

Table 1 K_i values of Y-27632-related compounds for protein serine-threonine kinases

Protein kinases	K_m for ATP (μ M)	K_i (μ M)				
		Y-27632	Y-30141	Y-30964	HA1077 ²⁷	H-7 ²⁸
p160ROCK	0.1	0.14	0.010	0.15	0.33	0.45
Protein kinase C (rat brain)	6	26	1.4	12	9.3	7.7
cAMP-dependent protein kinase	13	25	0.29	0.65	1.0	5.7
Myosin light-chain kinase	50	>250	14	19	55	170

Protein kinases were purified and assayed in the presence of various concentrations of the inhibitors, and the drug concentration required to inhibit the kinase activity by 50% (IC_{50} values) was obtained. K_i values were calculated according to the equation, $K_i = IC_{50}/(1 + S/K_m)$, where S and K_m are the concentration of and K_m value for ATP. The K_m values of cAMP-dependent protein kinase and myosin light-chain kinase are from refs 26 and 29, respectively.

Figure 2 Identification of p160ROCK as a target for Y-27632. **a**, Binding of [3 H]Y-30141 to aortic membranes. ●, specific binding; ○, total binding; ▲, nonspecific binding in the presence of 10 μ M Y-30141. **b**, Scatter plots of the IC_{50} values of 35 compounds for smooth muscle contraction and for [3 H]Y-30141 binding. 1, Y-33075; 2, Y-30141; 4, Y-32885; 8, Y-27632; 10, Y-30964; 16, Y-28791. 23, staurosporine; 24, K252a; 25, HA1077; 26, HA1004; 27, H-7; 28, ML-9. **c**, Photo-affinity labelling of the target protein. Irradiation was performed for 0, 1 and 5 min in the absence or presence of 1 mM Y-30141. **d**, Inhibition by Y-27632 and related compounds of p160ROCK kinase activity.

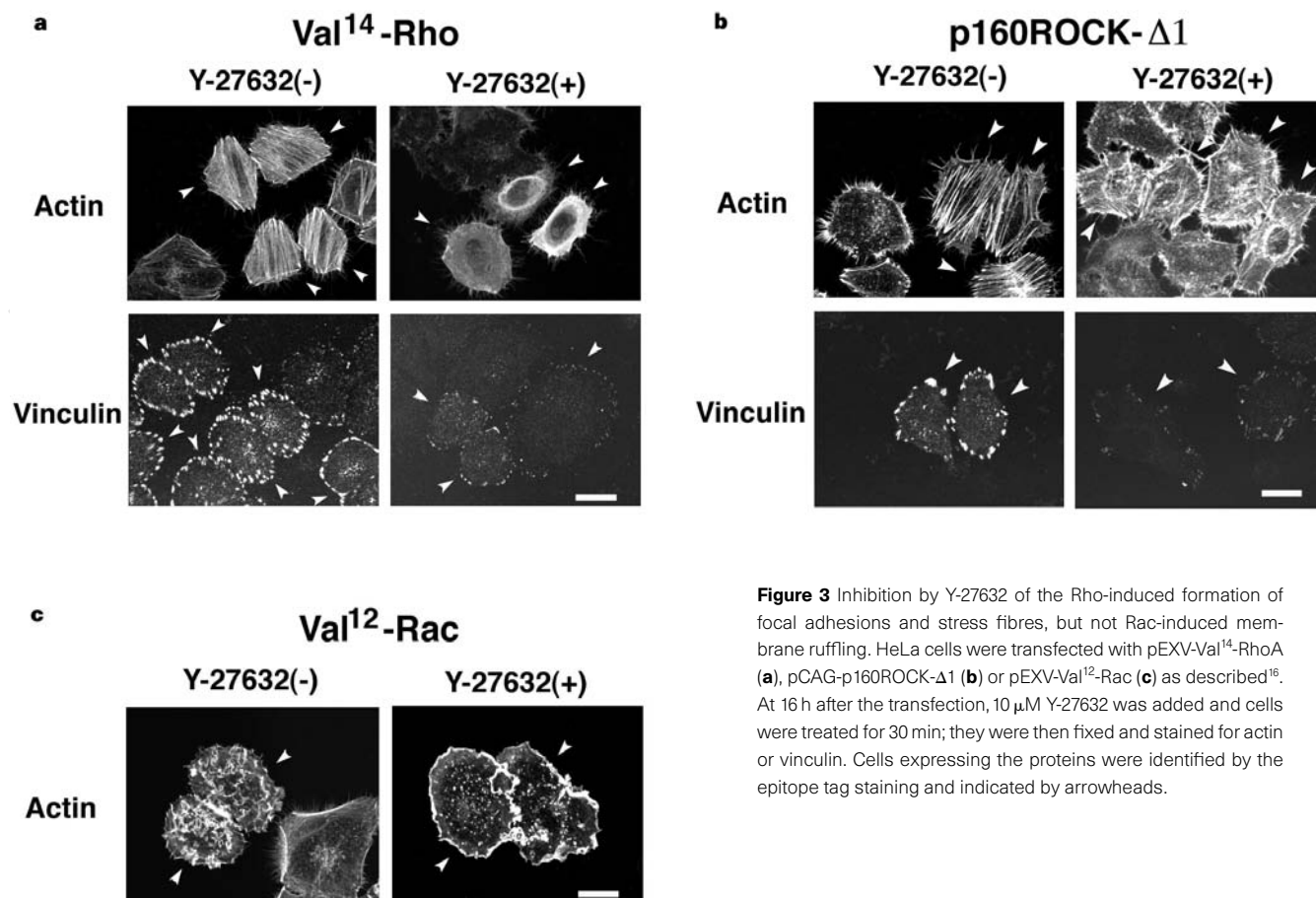
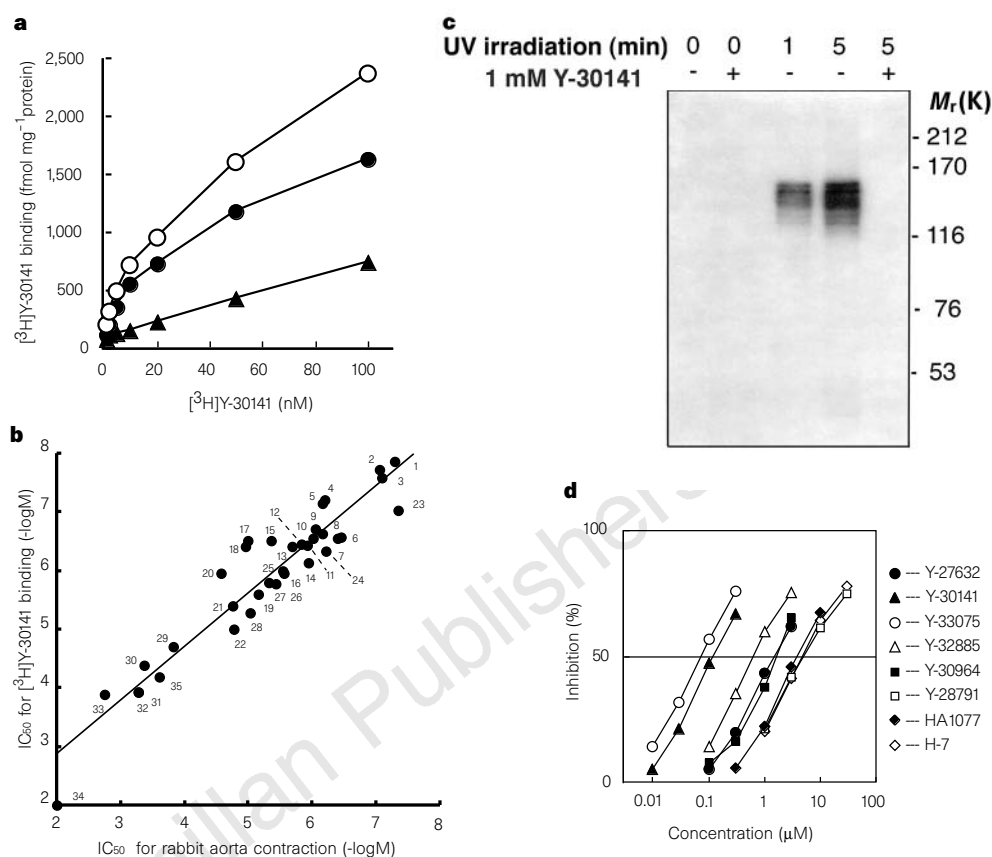


Figure 3 Inhibition by Y-27632 of the Rho-induced formation of focal adhesions and stress fibres, but not Rac-induced membrane ruffling. HeLa cells were transfected with pEXV-Val 14 -RhoA (**a**), pCAG-p160ROCK- Δ 1 (**b**) or pEXV-Val 12 -Rac (**c**) as described¹⁶. At 16 h after the transfection, 10 μ M Y-27632 was added and cells were treated for 30 min; they were then fixed and stained for actin or vinculin. Cells expressing the proteins were identified by the epitope tag staining and indicated by arrowheads.

pressure in a dose-dependent manner in spontaneous hypertensive rats: a fall of 50 mm Hg was still observed 7 h after administration of 30 mg kg⁻¹ of Y-27632. The same dose of this compound also caused a significant and persistent fall in blood pressure in renal hypertensive rats, as well as in deoxycorticosterone acetate (DOCA)-salt hypertensive rats (Fig. 4b, c). On the other hand, administration of the same dose of Y-27632 caused only a slight and transient fall in blood pressure in control Wistar rats (Fig. 4d). These results suggest that Rho/ROCK-mediated Ca²⁺ sensitization contributes to blood-pressure regulation *in vivo* and that it is augmented in hypertension.

Y-27632 is thus a valuable tool for investigating the functions of

p160ROCK *in vivo* and its pathophysiological implications. As well as modulating smooth-muscle contraction, the Rho/ROCK pathway may help to regulate integrin-mediated cell adhesion and motility, which could be critical in processes such as tumour-cell metastasis and immunoactivation^{19,21}. Y-27632 should be useful for investigating the role of p160ROCK in these processes and may be clinically important. □

Methods

Compounds. The synthesis of compounds used in this study will be reported in detail elsewhere. Y-27632, (R)-(+)-*trans*-N-(4-pyridyl)-4-(1-aminoethyl)-cyclohexanecarboxamide; Y-30141, (R)-(+)-*trans*-N-(1H-pyrrolo[2,3-*b*]pyridin-4-yl)-4-(1-aminoethyl)cyclohexanecarboxamide; Y-35526, (R)-(+)-N-(3-iodo-1H-pyrrolo[2,3-*b*]pyridin-4-yl)-4-(1-aminoethyl)-3-azidobenzamide; Y-28791, *trans*-N-(4-pyridyl)-4-guanidinomethylcyclohexanecarboxamide; Y-30964, *trans*-N-(1H-pyrazolo[3,4-*b*]pyridin-4-yl)-4-guanidinomethylcyclohexanecarboxamide; Y-32885, (R)-(+)-N-(4-pyridyl)-4-(1-aminoethyl)-benzamide; Y-33075, (R)-(+)-N-(1H-pyrrolo[2,3-*b*]pyridin-4-yl)-4-(1-aminoethyl)benzamide. [³H]Y-30141 was synthesized by a catalytic reduction of the corresponding aryl bromide with tritium gas. [¹²⁵I]Y-35526 was synthesized from the parent compound by the chloramine-T method.

Bioassay. Aortic rings were prepared from male rabbits weighing 1.9 to 3.0 kg and suspended as described²². After the ring was contracted with 50.9 mM KCl or 1 μM phenylephrine for 1 h, a test compound was added cumulatively. Pig coronary artery and guinea-pig trachea were prepared and the relaxant effects of the test compounds were also examined.

Rabbit mesenteric artery was permeabilized by incubating with 20 μM β-escin for 15 min²³. Contractile responses were evoked by addition of either CaCl₂, 10 μM GTPγS or 300 nM calyculin A. The free Ca²⁺ concentration was estimated from the binding constant of Ca²⁺ and EGTA, 10⁶ M⁻¹.

Radioligand binding assay with [³H]Y-30141. The bovine aorta microsome (135 μg protein) was incubated at 30 °C for 90 min with 10 nM [³H]Y-30141 (633 Bq μmol⁻¹) with or without test compounds in 50 mM Tris-HCl, pH 7.4, containing 5 mM MgCl₂ and 0.25% BSA in 0.2 ml. The reaction was terminated by filtration (Whatman GF/C filters) and the radioactivity on the filter counted.

Purification and peptide sequencing of the Y-27632-binding protein. Bovine aorta was minced, homogenized by a Polytron homogenizer in nine volumes of buffer A (50 mM Tris-HCl, pH 7.4, 1 mM dithiothreitol, 1 mM EGTA, 1 mM EDTA, and a mixture of protease inhibitors), and centrifuged at 1,000g for 30 min. The supernatant was centrifuged at 10,000g for 30 min. The supernatant was applied to a hydroxyapatite column equilibrated with 10 mM potassium phosphate, pH 6.8, containing 1 mM dithiothreitol, 100 μM (*p*-amidinophenyl)methylsulphonyl fluoride and 0.02% sodium azide. The column was washed with 0.2 M potassium phosphate (pH 6.8) and eluted with 0.4 M potassium phosphate (pH 6.8). To identify the Y-27632-binding protein, we incubated an aliquot of the eluate with 1.5 nM [¹²⁵I]Y-35526 (74 Bq μmol⁻¹) at 25 °C for 1 h in the dark. The mixture was then irradiated with 254-nm light (8 W) for 5 min at 3 cm. Labelled proteins were identified by SDS-PAGE and autoradiography. To purify the Y-27632-binding protein, the eluate was fractionated with ammonium sulphate at 40% saturation and the pellet dialysed against buffer A. The dialysate was applied to a Bio-Rad preparative SDS-PAGE system Model 491. Fractions containing the 150K protein were desalted using Centricon 10 and dried. Residues were dissolved in 70% formic acid. Solid BrCN was added and the mixture incubated at room temperature overnight. Cleaved peptides were separated by SDS-PAGE and transferred to a PVDF membrane. The principal peptides at 6.5K and 18K were sequenced.

Kinase assay. p160ROCK, ROCK-II and PAK were expressed in COS cells as tagged full-length proteins^{16,24}. Expressed proteins were immunoprecipitated by the use of anti-tag antibodies and assayed as described⁷. Protein kinase C was partially purified from rat brain. Protein kinase C and cAMP-dependent protein kinase (Sigma) were assayed according to ref. 25. Myosin light-chain kinase was prepared from chicken gizzard as described²⁶ without a DEAE-Sephacel step, and assayed using 10 μM ATP with chicken gizzard myosin (110 μg ml⁻¹) as substrate.

Rat hypertension models. Seven-week-old male Wistar rats were anaesthe-

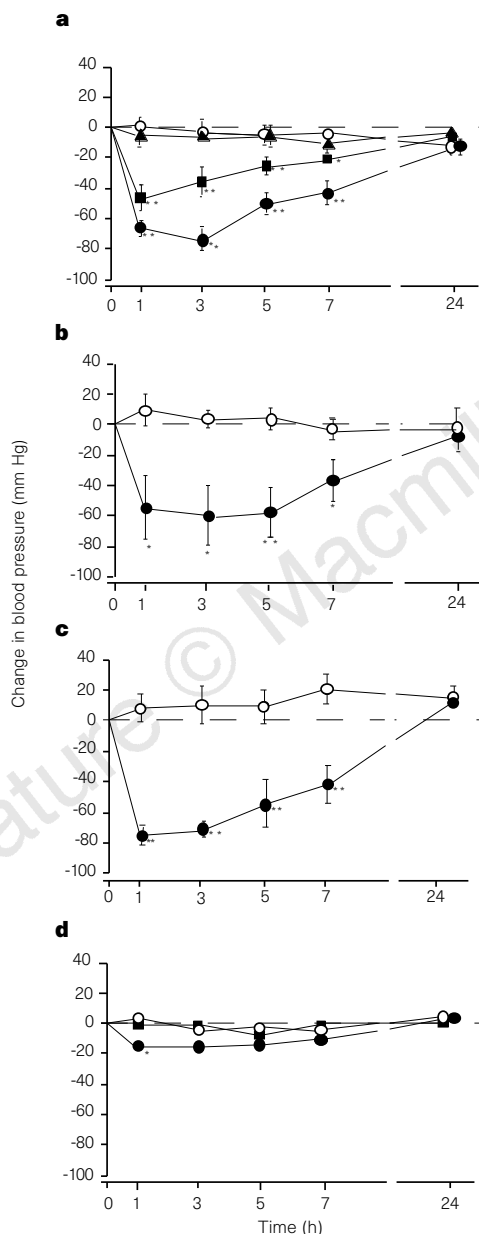


Figure 4 Effects of Y-27632 on blood pressure. Y-27632 was administered *per os* to spontaneously hypertensive rats (a), renal hypertensive rats (b), DOCA-salt rats (c) and control Wistar rats (d), and the systolic blood pressure was followed. O, vehicle alone; ▲, 3; ■, 10; and ●, 30 mg kg⁻¹ of Y-27632. Each value represents the mean ± s.e.m. of four to six experiments. Significant difference was evaluated by Student's unpaired *t*-test; **P* < 0.05 and ***P* < 0.01, compared to controls with vehicle alone.

sized with sodium pentobarbital. A silver clip (0.2 mm in diameter) was placed on the left renal artery in the preparation of the renal hypertensive rats. In the preparation of the DOCA-salt hypertensive rats, the left kidney was removed and a DOCA pellet (50 mg) was implanted subcutaneously. The DOCA rats were then fed an 8% salt diet. Rats from both groups were used after 8 weeks in the experiments, together with male, 17–22-week old spontaneous hypertensive rats. The average systolic pressure in these groups of hypertensive rats ranged from 209 to 237 mm Hg, and no significant difference was found between groups. Eight-week-old male Wistar rats were used as controls. Their average systolic pressure was 139 mm Hg. Y-27632 was administered orally. The systolic blood pressure was measured by the tail cuff method at 1, 3, 5, 7 and 24 h. The rats were prewarmed to 40 °C for 10 min before each measurement. No toxicity was found in rats treated with 30 mg kg⁻¹ of Y-27632 administered *per os* once per day for 10 days.

Received 3 June; accepted 13 August 1997.

- Somlyo, A. P. & Somlyo, A. V. Signal transduction and regulation in smooth muscle. *Nature* **372**, 231–236 (1994).
- Somlyo, A. P. *et al.* Modulation of Ca²⁺-sensitivity and of the time course of contraction in smooth muscle: A major role of protein phosphatase? *Adv. Prot. Phosphat.* **5**, 181–195 (1989).
- Kitazawa, T., Masuo, M. & Somlyo, A. P. G Protein-mediated inhibition of myosin light-chain phosphatase in vascular smooth muscle. *Proc. Natl Acad. Sci. USA* **88**, 9307–9310 (1991).
- Hirata, K. *et al.* Involvement of rho p21 in the GTP-enhanced calcium ion sensitivity of smooth muscle contraction. *J. Biol. Chem.* **267**, 8719–8722 (1992).
- Gong, M. C. *et al.* Role of guanine nucleotide-binding proteins, ras-family or trimeric proteins or both in Ca²⁺ sensitization of smooth muscle. *Proc. Natl Acad. Sci. USA* **93**, 1340–1345 (1996).
- Kimura, K. *et al.* Regulation of myosin phosphatase by Rho and Rho-associated kinase (Rho-kinase). *Science* **273**, 245–248 (1996).
- Ishizaki, T. *et al.* The small GTP-binding protein Rho binds to and activates a 160 kDa Ser/Thr protein kinase homologous to myotonic dystrophy kinase. *EMBO J.* **15**, 1885–1893 (1996).
- Kitazawa, T., Kobayashi, S., Horii, K., Somlyo, A. V. & Somlyo, A. P. Receptor-coupled, permeabilized smooth muscle. *J. Biol. Chem.* **264**, 5339–5342 (1989).
- Ishihara, H. *et al.* Calyculin A and okadaic acid: inhibitors of protein phosphatase activity. *Biochem. Biophys. Res. Commun.* **159**, 871–877 (1989).
- Saitoh, M., Ishikawa, T., Matsumura, S., Naka, M. & Hidaka, H. Selective inhibition of catalytic activity of smooth muscle myosin light chain kinase. *J. Biol. Chem.* **262**, 7796–7801 (1987).
- Leung, T., Manser, E., Tan, L. & Lim, L. A novel serine/threonine kinase binding the Ras-related RhoA GTPase which translocates the kinase to peripheral membranes. *J. Biol. Chem.* **270**, 29051–29054 (1995).
- Matsui, T. *et al.* Rho-associated kinase, a novel serine threonine kinase, as a putative target for the small GTP binding protein Rho. *EMBO J.* **15**, 2208–2216 (1996).
- Nakagawa, O. *et al.* ROCK-I and ROCK-II; two isoforms of Rho-associated coiled-coil forming protein serine/threonine kinase in mice. *FEBS Lett.* **392**, 189–193 (1996).
- Manser, E. *et al.* A brain serine/threonine protein kinase activated by Cdc42 and Rac1. *Nature* **367**, 40–46 (1994).
- Ridley, A. J. & Hall, A. The small GTP-binding protein rho regulates the assembly of focal adhesions and actin stress fibers in response to growth factors. *Cell* **70**, 389–399 (1992).
- Ishizaki, T. *et al.* p160ROCK, a Rho-associated coiled-coil forming protein kinase, works downstream of Rho and induces focal adhesions. *FEBS Lett.* **404**, 118–124 (1997).
- Leung, T., Chen, X.-Q., Manser, E. & Lim, L. The p160 RhoA-binding kinase ROK α is a member of a kinase family and is involved in the reorganization of the cytoskeleton. *Mol. Cell. Biol.* **16**, 5313–5327 (1996).
- Amano, M. *et al.* Formation of actin stress fibres and focal adhesions enhanced by Rho-kinase. *Science* **275**, 1308–1311 (1997).
- Narumiya, S. The small GTPase Rho: cellular functions and signal transduction. *J. Biochem.* **120**, 215–228 (1996).
- Ridley, A. J., Paterson, H. F., Johnston, C. L., Diekmann, D. & Hall, A. The small GTP-binding protein rac regulates growth factor-induced membrane ruffling. *Cell* **70**, 401–410 (1992).
- Tapon, N. & Hall, A. Rho, Rac and Cdc42 GTPases regulate the organization of the actin cytoskeleton. *Curr. Opin. Cell Biol.* **9**, 86–92 (1997).
- Sato, H. & Inui, J. Endothelial cell-dependent relaxation and contraction induced by histamine in the isolated guinea-pig pulmonary artery. *Eur. J. Pharmacol.* **97**, 321–324 (1984).
- Kobayashi, S., Kitazawa, T., Somlyo, A. V. & Somlyo, A. P. Cytosolic heparin inhibits muscarinic and α -adrenergic Ca²⁺ release in smooth muscle. *J. Biol. Chem.* **264**, 17997–18004 (1989).
- Bagrodia, S. *et al.* Identification of a mouse p21^{Cdc42/Rac} activated kinase. *J. Biol. Chem.* **270**, 22731–22737 (1995).
- Kikkawa, U., Takai, Y., Minakuchi, R., Inohara, S. & Nishizuka, Y. calcium-activated, phospholipid-dependent protein kinase from rat brain. Subcellular distribution, purification and properties. *J. Biol. Chem.* **257**, 13341–13348 (1982).
- Adelstein, R. S. & Klee, C. B. Purification of smooth muscle myosin light chain kinase. *Meth. Enzym.* **85**, 298–308 (1985).
- Takayasu, M. *et al.* The effects of HA compound calcium antagonists on delayed cerebral vasospasm in dogs. *J. Neurosurg.* **65**, 80–85 (1986).
- Hidaka, H., Inagaki, M., Kawamoto, S. & Sasaki, Y. Isoquinolinesulfonamides, novel and potent inhibitors of cyclic nucleotide dependent protein kinase and protein kinase C. *Biochemistry* **9**, 5036–5041 (1984).
- Bubin, C. S., Erlichman, J. & Rosen, O. M. Cyclic AMP-dependent protein kinase from bovine heart muscle. *Meth. Enzymol.* **38**, 308–315 (1974).

Acknowledgements. We thank S. Ohno and R. Cerione for recombinant PKC ϵ and PAK cDNA, respectively, H. Karaki for helpful suggestions, H. Bito for critical reading of the manuscript, and K. Okuyama for secretarial assistance. Supported in part by a grant-in-aid for specially promoted research from the Ministry of Education, Science and Culture of Japan, and by an HFSP grant.

Correspondence and requests for materials should be addressed to M.U. (e-mail: uehata@yoshitomi.co.jp) or S.N. (e-mail: snaru@four.med.kyoto-u.ac.jp).

The Fork head transcription factor DAF-16 transduces insulin-like metabolic and longevity signals in *C. elegans*

Scott Ogg, Suzanne Paradis, Shoshanna Gottlieb, Garth I. Patterson, Linda Lee, Heidi A. Tissenbaum & Gary Ruvkun

Department of Molecular Biology, Massachusetts General Hospital, Department of Genetics, Harvard Medical School, Boston, Massachusetts 02114, USA

In mammals, insulin signalling regulates glucose transport together with the expression and activity of various metabolic enzymes. In the nematode *Caenorhabditis elegans*, a related pathway regulates metabolism, development and longevity^{1,2}. Wild-type animals enter the developmentally arrested dauer stage in response to high levels of a secreted pheromone³, accumulating large amounts of fat in their intestines and hypodermis. Mutants in DAF-2 (a homologue of the mammalian insulin receptor) and AGE-1 (a homologue of the catalytic subunit of mammalian phosphatidylinositol 3-OH kinase) arrest development at the dauer stage³. Moreover, animals bearing weak or temperature-sensitive mutations in *daf-2* and *age-1* can develop reproductively, but nevertheless show increased energy storage and longevity^{1,2,4,5}. Here we show that null mutations in *daf-16* suppress the effects of mutations in *daf-2* or *age-1*; lack of *daf-16* bypasses the need for this insulin receptor-like signalling pathway. The principal role of DAF-2/AGE-1 signalling is thus to antagonize DAF-16. *daf-16* is widely expressed and encodes three members of the Fork head family of transcription factors. The DAF-2 pathway acts synergistically with the pathway activated by a nematode TGF- β -type signal, DAF-7, suggesting that DAF-16 cooperates with nematode SMAD proteins in regulating the transcription of key metabolic and developmental control genes. The probable human orthologues of DAF-16, FKHR and AFX, may also act downstream of insulin signalling and cooperate with TGF- β effectors in mediating metabolic regulation. These genes may be dysregulated in diabetes.

The metabolic, longevity, and developmental defects caused by *daf-2* and *age-1* mutations are suppressed by *daf-16* mutations^{5,6} (Fig. 1 and Table 1). Wild-type animals do not arrest at the dauer stage in low pheromone, whereas temperature-sensitive mutations in *daf-2* cause 100% arrest at the dauer larval stage (Table 1). *daf-2* mutant animals also shift metabolism towards fat accumulation and show large drops of fat in the intestine and hypodermis of dauer larvae, as measured by Sudan black staining (Fig. 1a). However, animals bearing both a *daf-2* mutation and the *daf-16(mgDf47)* or *daf-16(mgDf50)* null mutations (see below) do not arrest at the dauer stage or store large amounts of fat, developing instead into reproductive adults (Table 1 and Fig. 1b). Therefore, *daf-16* is a major target of DAF-2/AGE-1 signalling.

A parallel DAF-7 TGF- β pathway also regulates *C. elegans* metabolism and development⁷. *daf-16* gene activity is not necessary for the metabolic shift towards energy storage or dauer arrest induced by lack of DAF-7 TGF- β signalling (Fig. 1c). For example, like *daf-7(e1372)* mutant animals, *daf-16(mgDf50)*; *daf-7(e1372)* mutant animals arrest at the dauer stage (172/174 animals arrest) and accumulate fat. The arrest, however, is not complete. The pharynx is not contracted (2/100 arrested animals have a constricted pharynx) and the intestine is less refractile than normal dauers, as has been reported using the *daf-16(m26)* allele⁷ (see below). Thus

Table 1 Effect of *daf-16* null mutations on dauer formation

Genotype	Phenotype of progeny at 25°C (%)				
	L4 and adults	Dauer	Dead eggs	Other*	N†
N2 Bristol (wild type)	98.6	0	0.3	1.1	721
<i>daf-2(e1370)</i>	0	95.2	2.8	2.0	714
<i>daf-16(mgDf50)</i>	79.6	0	18.1	2.3	348
<i>daf-16(mgDf47); daf-2(e1370)</i>	84.1	0	10.2	5.7	422
<i>daf-16(mg50); daf-2(e1370)</i>	95.1	0	3.5	1.4	833
<i>daf-16(mg54); daf-2(e1370)</i>	89.6	0.4‡	9.2	0.8	502
<i>daf-3(mgDf90)</i>	98.6	0	ND	1.4	737
<i>daf-2(e1370); daf-3(mgDf90)</i>	0	97	ND	3.0	934
Pheromone + wild type	64.2	35.8	ND	0	1,121
Pheromone + <i>daf-16(mgDf50)</i>	73.9	23.9‡	ND	2.2	742
Pheromone + <i>daf-3(mgDf90)</i>	61.0	37.7	ND	1.3	899
Pheromone + <i>daf-16(mgDf50); daf-3(mgDf90)</i>	56.1	41.8‡	ND	2.1	743

* 'Other' includes animals that could not be classified as dauer or non-dauer because the animal was young, had grossly aberrant morphology, or was dead.

† N, total number of animals scored.

‡ Under these conditions the dauers were not fully formed. These larvae were similar to dauers in having a shrunken cuticle, but the pharynx pumped and was not shrunken, the intestine was not dark and alae were indistinct.

placement of *daf-16* in the insulin receptor-like genetic pathway and not the DAF-7 TGF- β -like pathway is confirmed with a null *daf-16* allele. In support of the parallel pathway model, the *daf-3(mgDf90)* null mutation, which potently suppresses *daf-7(e1372)* (ref. 8), does not suppress *daf-2(e1370)* (Table 1 and Fig. 1d).

daf-16 and *daf-3* do not represent the only output of pheromone signalling. *daf-16(mgDf50)* animals arrest at the dauer stage in dauer pheromone, although the arrested animals are partial dauers, with large pharynxes and low refractility (Table 1; see below). *daf-3(mgDf90)* single mutants form apparently normal dauers in the presence of pheromone (Table 1). *daf-16; daf-3* double-null mutant animals arrest as partial dauers in dauer pheromone, like the *daf-16* single mutant (Table 1). Thus, *daf-16* gene activity is necessary for some aspects of dauer arrest, and other pathways besides those coupled to *daf-3* and *daf-16*, for example the pathways defined by *daf-11* or *daf-12* (ref. 3), may also regulate arrest at the dauer stage in pheromone. *daf-16* and *daf-3* have no obvious function during reproductive development. Reproductively growing animals bearing null mutations in either *daf-16* or *daf-3*, or mutations in both genes do not dramatically alter fat

accumulation in larvae or adult animals compared to wild type (data not shown).

We genetically mapped *daf-16* to a 300-kilobase cosmid and YAC contig (Fig. 2a). Multiple *daf-16* mutations induced with γ rays or ethylmethane sulphonate (EMS) were isolated as suppressors of the dauer constitutive phenotype induced by the *daf-2(e1370)* mutant. DNAs isolated from these *daf-16* alleles were surveyed for deletions or rearrangement breakpoints using cosmids in this contig (Fig. 2a). Deletion and point mutations associated with *daf-16* alleles (see below) were localized within cosmid R13H8 to a gene that encodes a member of the Fork head family of transcription factors. The *daf-16(mgDf50)* null allele (see below) was rescued by injection of cosmid R13H8 into the germ line of a *daf-16(mgDf50); daf-2(e1370)* double mutant and showing that transgenic progeny arrest at the dauer stage. A neighbouring cosmid did not rescue this *daf-16* mutant allele.

Analysis of *daf-16* cDNAs revealed three alternatively spliced forms, designated *daf-16a1*, *daf-16a2* and *daf-16b* (Fig. 2a). DAF-16a and DAF-16b have distinct, but highly related, Fork head type DNA-binding domains (73% identical (87/119 amino acids); Fig. 2c). The amino-terminal half of the DNA-binding domain is encoded by exons 3 and 4 for DAF-16a and exon 5 for DAF-16b. The C-terminal half of each DNA-binding domain is encoded by exons 6 and 7 which are common to both transcripts (Fig. 2). Two similar isoforms of DAF-16a, DAF-16a1 and DAF-16a2 are encoded by the same 11 exons, except that exon 3 is alternatively spliced, resulting in the addition of two amino acids to DAF-16a1 outside the DNA-binding domain (Fig. 2a, b).

The DAF-16a and DAF-16b isoforms are probably generated from distinct promoters, suggesting that they could be expressed in and mediate DAF-2 signalling in distinct tissues or cell types (see below). These isoforms have distinct Fork head DNA-binding domains that may interact with distinct partners (for example, DAF-3, DAF-8 or DAF-14; see below) or may bind to distinct downstream promoters. Significantly, the probable DNA-binding specificity determinant of Fork head proteins in helix 3 (ref. 3) is common to the DAF-16 isoforms. Thus, the expression and differential splicing of the two distinct DAF-16 Fork head DNA-binding domains is expected to be functionally important but is not obviously implicated in DNA-binding specificity.

The most closely related proteins in GenBank to DAF-16 are human FKHR and AFX (Fig. 2c, d). Within the Fork head DNA-binding domain, DAF-16a is 65% identical (78/120) to FKHR and 62% identical (74/120) to AFX. DAF-16b is 50% identical (60/120) to FKHR and 47% identical (56/120) to AFX. Within this region, FKHR and AFX are 81% identical (95/117). In addition, there is a region at the amino-terminal end of each protein that is 55% identical (10/18) between DAF-16a and FKHR and AFX. This region is not conserved between other Fork head-related proteins

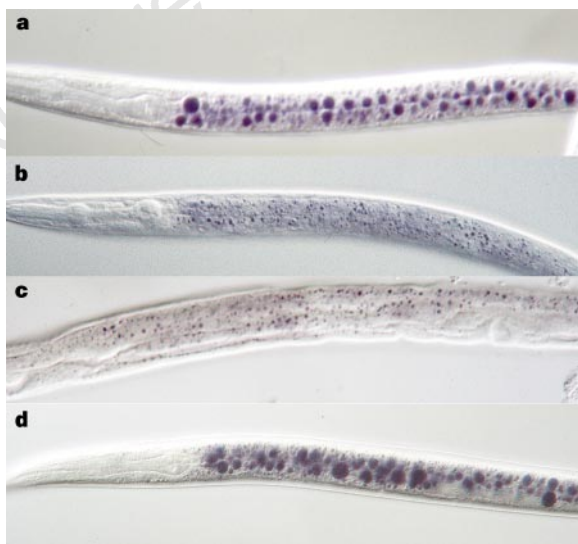


Figure 1 Metabolic control by *daf-16* and *daf-3*. Fat accumulation assayed by Sudan black in dauer larval stage and the comparable reproductive larval stage 3 and 4 in animals grown at 25°C. **a**, *daf-2(e1370)*. **b**, Suppression of fat accumulation in *daf-16(mgDf50); daf-2(e1370)* animals. **c**, Suppression of *daf-7(e1372)* fat storage and arrest by *daf-3(mgDf90)*. **d**, *daf-16(mgDf50)* mutant does not suppress the fat accumulation or arrest of a *daf-7(e1372)* mutant.

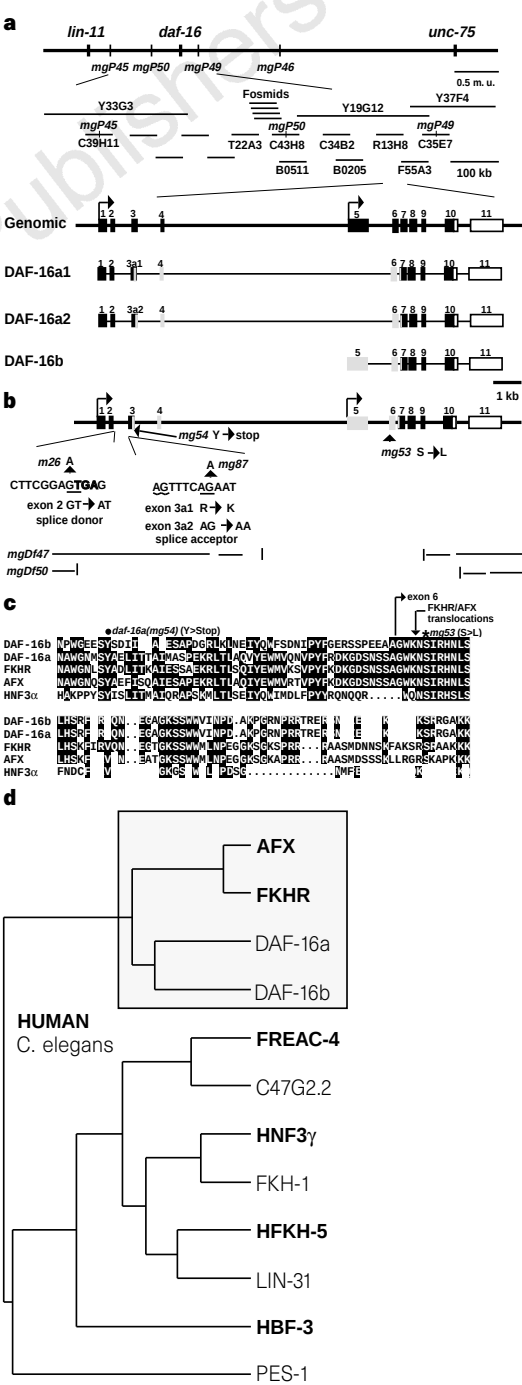
nor between DAF-16b and FKHR or AFX. DAF-16a is only 36% identical to Freac-4 and DAF-16b is only 33% identical to Freac-3 (Fig. 2d), which are the next most closely related proteins in the GenBank or dbEST databases. DAF-16a and DAF-16b are similarly distinct from the closest *C. elegans* homologues (Fig. 2d). Phylogenetic tree analysis of the Fork head domain of DAF-16 and its closest relatives show that FKHR, AFX and DAF-16 constitute a distinct class (Fig. 2d).

The high degree of sequence similarity suggests that FKHR and AFX may correspond to the human orthologues of DAF-16 and constitute major transcriptional outputs of insulin or IGF-1 receptors in mammals. FKHR and AFX, however, were identified as oncogene breakpoints and not, like DAF-16, as metabolic transcriptional regulators. The oncogenic translocations fuse a disrupted FKHR or AFX Fork head DNA-binding domain and intact carboxy-

terminal activation domain to other DNA-binding domains^{10,11}. The PAX-3/FKHR fusion protein is a more potent transcriptional activator than PAX-3 (ref. 12). Although FKHR and AFX activities are not yet known to be regulated by insulin or IGF, growth of rhabdomyosarcomas is suppressed by declines in IGF-1 receptor signalling^{13,14} and these tumours accumulate large amounts of glycogen¹⁵. The HNF3 Fork head transcription factor has been implicated in insulin regulation of metabolic gene transcription¹⁶. Because distantly related Fork head proteins bind to similar sites¹⁷, HNF3-binding *cis*-acting sites from insulin-responsive metabolic control genes may actually be regulated by the FKHR or AFX protein in mammals. Non-null mutations in FKHR or AFX that decouple these transcription factors from the insulin signalling cascade may underlie type II diabetes in some pedigrees.

Identification of *daf-16* null alleles establishes that the most

Figure 2 *daf-16* encodes a Fork head/HNF3-related transcription factor. **a**, Top, genetic and physical map of the *daf-16* region, bottom, exon/intron structure of *daf-16*. Coding regions are filled boxes, non-coding regions are open boxes, and introns are lines. The Fork head regions are indicated in grey. **b**, Molecular identity of six mutations identified in *daf-16*. The genomic region of *daf-16* is shown. Boxes indicate exons, and areas in grey indicate regions homologous to the Fork head domain. Two point mutants (*mg54* and *mg53*) are shown simply as arrows pointing to the location of the lesion (see also **c**). Two point mutants (*m26* and *mg87*) affect splice sites and are shown below. The extent of the *mgDf47* and *mgDf50* deletions is indicated. The area between the vertical lines is deleted, and areas of uncertainty are indicated by dotted lines. **c**, DNA-binding domains of DAF-16a, DAF-16b and human FKHR (U02310), AFX (X93996) and HNF3 α were aligned using Pileup (GCG). Amino acids that are identical in at least three of the proteins are highlighted. Horizontal arrow indicates the beginning of exon 6 which is shared between DAF-16a and DAF-16b. Point mutations within the DNA-binding domains of DAF-16 are indicated; *daf-16(mg54)* is specific for DAF-16a, whereas *daf-16(mg53)* is common for both DAF-16a and DAF-16b. Vertical arrow indicates the fusion point for the PAX-3/FKHR and HTRX1/AFX-1 fusion proteins that are caused by chromosomal translocations found in human tumours. **d**, Pileup (GCG) was used to align the DNA binding domains of the listed proteins. Human proteins are indicated in bold text and *C. elegans* proteins are indicated by plain text. C47G2.2 is a gene-finder-predicted ORF on cosmid C47G2. The accession numbers for the *daf-16* sequences are AF020342 *daf-16a1*; AF020343 *daf-16a2* and AF020344 *daf-16b*.



Genotype	% Dauer formation		
	15°C	18°C	20°C
<i>daf-2(e1370)</i>	0 (n = 611)	5 (n = 487)	5 (n = 506)
<i>daf-1(m40)</i>	0 (n = 434)	0 (n = 732)	1 (n = 909)
<i>daf-2(e1370); daf-1(m40)</i>	17 (n = 652)	99 (n = 861)	100 (n = 718)

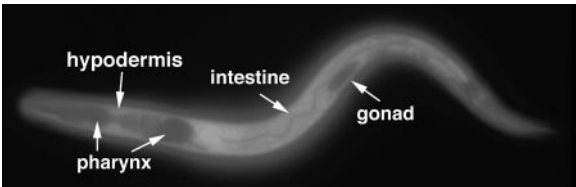


Figure 3 Expression of a *daf-16a::GFP* fusion gene in an L1 animal.

important *daf-16* function is in dauer arrest and metabolic regulation. *daf-16(mgDf47)* and *daf-16(mgDf50)* were isolated after γ -ray mutagenesis and are molecular null alleles that delete conserved domains or nearly all of the *daf-16* coding region (Fig. 2b). As already described, both of these *daf-16* null alleles have the same phenotypes: viable and complete suppression of the *daf-2* dauer constitutive and energy storage phenotypes.

Molecular analysis of other *daf-16* mutant alleles revealed that the two major DAF-16 isoforms are not redundant. *daf-16(mg54)* is an amber stop mutation within the DNA-binding domain of DAF-16a but is predicted not to affect DAF-16b (Fig. 2b, c). *daf-16(mg54)* suppresses the dauer constitutive phenotype caused by *daf-2* mutations (Table 1). Two other *daf-16* alleles, *m26* and *mg87*, are also specific to the *daf-16a* isoform (Fig. 2c). *daf-16(m26)* incompletely suppresses *daf-2(e1370)* (ref. 7). The identification of *daf-16a*-specific mutations indicates that *daf-16b* activity is not sufficient to induce dauer arrest in a *daf-2* mutant. No mutation that is specific for the *daf-16b* Fork head region was detected in a collection of 11 *daf-16* alleles sequenced, so it is not known if this isoform is necessary for *daf-2* and *age-1* induced dauer arrest. The only point mutation that affects both *daf16* isoforms, *daf-16(mg53)*, changes a conserved serine in the Fork head DNA-binding domain to leucine (Fig. 2b, c).

Arrest at the dauer larval stage affects the morphology and metabolism of many *C. elegans* tissues³. DAF-16 might act directly in these target tissues or indirectly in neuronal or other signalling cells that in turn regulate the development and metabolism of target tissues. *daf-16* expression was addressed using a *daf-16a::green fluorescent protein (GFP)* fusion gene. Transgenic *daf-16a::GFP* animals show strong GFP fluorescence in most cells including ectoderm, muscles, intestine and neurons (Fig. 3). The broad expression pattern of *daf-16a::GFP* is also seen in late embryos, larvae and dauer larvae. No expression was seen in the pharynx even though the morphology of the pharynx is altered in dauer larvae, and *daf-16* gene activity is essential for this pharyngeal change in pheromone (Table 1) or *daf-7*-induced dauers⁷. It is possible that *daf-16b* is expressed in the pharynx and serves this function, that the *daf-16a::GFP* fusion gene does not bear those regulatory sequences, or that *daf-16* acts non-autonomously to affect the pharynx.

The expression of *daf-16a* in the many target tissues that are metabolically and developmentally shifted in dauers is consistent with its action in target tissues to regulate dauer arrest, metabolism and longevity. In addition to the metabolic changes that occur in dauer larvae, some tissues are remodelled or developmentally arrested. The hypodermis shrinks and stores fat. The intestinal lumen is shrunken and becomes refractile. The collagen composition of the cuticle is altered by changes in the expression of collagen genes³. *daf-16* is expressed in all of these tissues, suggesting that these phenotypic changes, including the metabolic shifts, may be induced by DAF-16-responsive genes.

daf-2 and *age-1* mutations also cause an increase in lifespan^{2,4,5,18}.

These effects are suppressed by *daf-16* mutations^{4,5,18}. Thus DAF-16 might directly regulate the genes necessary for increased longevity. For example, dauer larvae express higher superoxide dismutase levels, which might protect them from oxygen radicals¹⁹. This gene may be directly regulated by DAF-16. The suppression of the longevity increases of *daf-2* and *age-1* mutants by *daf-16* mutations suggests that other outputs of the DAF-2/AGE-1 signalling pathway, for example the regulation of glucose transport or metabolic enzyme activities, are not relevant to lifespan or dauer arrest in the absence of DAF-16 transcriptional regulation.

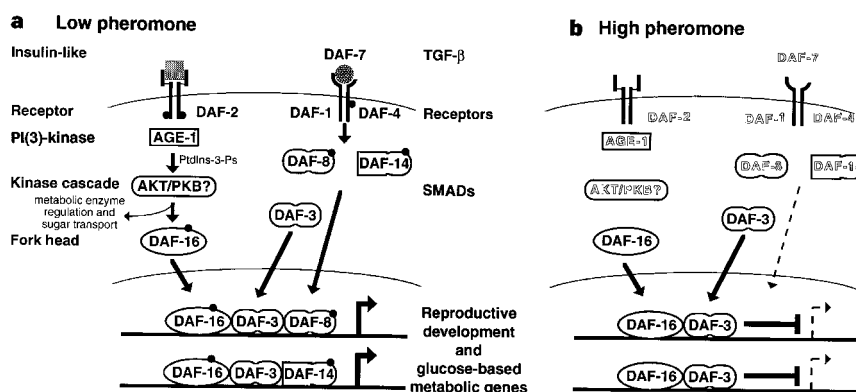
In extensive genetic screens for *daf* mutants, the only suppressors of *daf-2* that have been identified are *daf-16* (many alleles)^{6,20} and *daf-18* (one allele)²⁰. Therefore, *daf-16* is likely to represent the principal negatively regulated output of this insulin receptor-like signalling pathway. The lack of other *daf-2* suppressors argues against the existence of another entire ligand/receptor signalling pathway between *daf-2* and *daf-16* and suggests that *daf-16* acts in the same cell as *daf-2*.

The finding that DAF-16 encodes proteins highly related to the Fork head class of transcription factors suggests a molecular model for the negative regulation of DAF-16 by upstream DAF-2/AGE-1 signalling (Fig. 4). DAF-2 insulin receptor-like engagement by an unidentified ligand activates the AGE-1 phosphatidylinositol-3-OH kinase (PI(3)K). In mammals, PI(3)K signalling activates protein kinase cascades, including the AKT/PKB kinase²¹ that regulates metabolic enzyme activities as well as glucose transport²². We suggest that the homologous DAF-2/AGE-1 regulated kinase cascade couples to the DAF-16 transcription factor (as well as to metabolic enzymes and transporters). We suggest that in the absence of DAF-2/AGE-1 signalling inputs, DAF-16 acts as a repressor of metabolic genes that mediate energy usage (Fig. 4b). The repression of these genes causes a metabolic shift to energy storage that can be relieved by loss of *daf-16* gene activity. When the DAF-2/AGE-1 signalling cascade is activated under reproductive growth conditions, we suggest that a kinase cascade leads to the modification of the transcriptional activity of DAF-16 (Fig. 4a). This signalling may either change DAF-16 to an activator (Fig. 4a) or inactivate DAF-16 repressor function to now allow the expression of metabolic genes necessary for reproductive development.

There is precedent for the model that DAF-16 transcriptional activity is modulated by upstream tyrosine kinase signals: in *C. elegans* vulval cell lineages, the activity of the LET-23 EGF receptor kinase is coupled to the activity of the LIN-31 Fork head transcription factor (P. Tan and S. Kim, personal communication). And similar to the suppression of *daf-2* mutant phenotypes by *daf-16* null mutations, *lin-31* mutations bypass the normal requirement of upstream LET-23 EGF-like receptor signalling for the execution of particular vulval cell lineages²³.

In addition to the DAF-2/DAF-16 pathway, the activity of the parallel DAF-7 TGF- β pathway also regulates *C. elegans* metabolism. We find genetic synergy between the DAF-2 and DAF-7

Figure 4 Model for regulation of dauer formation by convergent transcriptional outputs from insulin receptor-like and TGF- β -like signal transduction pathways. **a**, Reproductive development occurs in low pheromone as a result of high DAF-2 ligand and DAF-7 signals. **b**, Dauer formation occurs when high pheromone levels downregulate DAF-2 ligand and DAF-7 signals.



signalling pathways by comparing the phenotype of a *daf-2; daf-1* double mutant to the phenotype of each single mutant (Table 2). This synergy, together with the similar phenotypes of mutants in the two pathways, implies that the pathways cooperate to control reproductive versus dauer development and metabolism. DAF-7 TGF- β is expressed by the exposed sensory neuron ASI under conditions that induce reproductive development but not in dauer pheromone.²⁴ Mutations that decrease or eliminate DAF-7 activity or the activities of the type I receptor DAF-1 (ref. 25) and type II receptor DAF-4 (ref. 26), or probable downstream SMAD proteins DAF-8 (A. Estevez and D. L. Riddle, personal communication), and DAF-14 (T. Inoue and J. Thomas, personal communication) cause arrest at the dauer stage and a metabolic shift to fat storage (Fig. 2). The dauer constitutive and fat storage phenotypes caused by mutations in these upstream TGF- β signalling genes are suppressed by null mutations in *daf-3*, which also encodes a SMAD protein⁸ (Fig. 2). Thus *daf-3* acts in this TGF- β pathway analogously to *daf-16* in the insulin-like pathway: *daf-3* gene activity is negatively regulated by upstream TGF- β signalling.

Similar convergence between tyrosine kinase receptor-mediated and TGF- β -related signalling pathways has been noted in early vertebrate development²⁷. Precedent from the regulation of *Xenopus* early development by activin/Smad2 signalling suggests a molecular model for convergent insulin receptor tyrosine kinase and TGF- β receptor regulation of development and metabolism: the Fork head protein FAST1 directly interacts with Smad2 cooperatively to bind to an activin-response element on a gene that is transcriptionally regulated by activin²⁸. The synergy between mutations in the *daf-2* insulin receptor-like and *daf-1* TGF- β receptor suggests that insulin-like transducing DAF-16 may interact with the TGF- β signal transducing SMAD proteins DAF-3, DAF-8 and/or DAF-14 on the promoters of genes that regulate metabolism and reproductive versus dauer development (Fig. 4).

We suggest that under low-pheromone conditions, when both DAF-2 ligand and DAF-7 endocrine signals are produced, DAF-16 activity is modified by an upstream DAF-2/AGE-1 kinase cascade, and DAF-8 and DAF-14 are phosphorylated by upstream DAF-7/DAF-1/DAF-4 receptor kinase signalling to induce a heteromeric Fork head/SMAD transcription factor complex (Fig. 4a). The complex activates expression of genes that mediate reproductive growth and metabolism. Under dauer-inducing conditions, we suggest that a low level of DAF-2 ligand does not activate phosphorylation of DAF-16 by the DAF-2/AGE-1 signalling cascade, and a low level of DAF-7 TGF- β does not activate phosphorylation of DAF-8 and DAF-14 by the DAF-1/DAF-4 receptor kinases. In these dauer-inducing conditions, we suggest that DAF-16 and DAF-3 form heteromers that repress the expression of reproductive development and metabolic genes.

Because *daf-3* (ref. 8), *daf-4* (ref. 8) and *daf-16* GFP fusion genes are widely expressed, we suggest that the integration of DAF-2 insulin receptor-like and DAF-7 TGF- β signalling takes place in the target tissues where the metabolic shifts are noted (Figs 1 and 4). The genes regulated by the Fork head/SMAD transcriptional regulatory complex may correspond to the *C. elegans* homologues of mammalian genes, such as those that encode PEPCK and the Glut4 glucose transporter, which are transcriptionally regulated by insulin²⁹. However, the molecular model does not depend on the signal integration taking place in target tissues. The signals may converge only in key regulatory cells, which in turn generate signals that regulate target tissue metabolism and development. In such a case, the downstream genes whose transcription is regulated by DAF-16 and DAF-3, DAF-8 and DAF-14 would not be target-tissue metabolic-control genes analogous to those regulated by insulin, but rather genes that encode other endocrine hormones.

It may be significant to the development of treatments for diabetes that *C. elegans* carrying mutations in *daf-16* can grow reproductively even if they also carry *daf-2* mutations that disable

insulin-receptor-like metabolic-control signals (Table 1). Thus the genetics yield the surprising result that these animals can live normally without insulin-receptor-like signalling if they are also mutant in *daf-16*. Some of the human metabolic defects that result from declines in insulin signalling in both type I and type II diabetes may be caused by unregulated activity of the human DAF-16 orthologues FKHR and AFX, in turn to cause changes in the expression of metabolic control enzymes. Inhibition of FKHR or AFX activities, for example by pharmaceutical agents that inhibit DNA binding or association with transcriptional partners, may ameliorate the metabolic defects caused by declines in insulin signalling. The more acute insulin responses, for example in glucose uptake and in metabolic enzyme activities, would not be directly affected by these transcriptional regulators.

The molecular model for convergent insulin-like and TGF- β signalling suggests that these pathways meet on the promoters of metabolic-control genes. Given such a mechanistic connection between the *C. elegans* insulin-like and TGF- β transcriptional outputs, this molecular complex may be conserved across phylogeny. We propose that, as in *C. elegans*, both insulin and a DAF-7-like neuroendocrine signal may be necessary for metabolic control in humans. Accordingly, the failure of target tissues to respond to insulin signals in type II diabetic patients could be due to defects in either the insulin or TGF- β signalling pathways. Genetic predisposition to type II diabetes has been noted. Declines in signalling from the human orthologues of DAF-7 (or in any of the signalling molecules downstream of DAF-7), could underlie the lack of response to insulin in type II diabetes, just as lack of DAF-7 signalling in *C. elegans* causes very similar metabolic defects as lack of DAF-2 signalling. Obesity is a major risk factor in type II diabetes. One attractive model for the connection between obesity and insulin signalling defects is that metabolic signals analogous to the dauer pheromone, which may be a fatty acid³⁰, or hormonal signals from fat cells may downregulate production of human DAF-7. □

Methods

***daf-16* molecular genetics.** DNA segments from the gap in the cosmid coverage of the *daf-16* genetic region were identified by probing a fosmid library with the end of cosmid T22A3, which defined the left endpoint of the gap that contains *daf-16*. One end of each fosmid so identified was identical to nucleotide sequence of two overlapping cosmids (C43H8 and B0511) near the end of a previously unlocalized 300 kb contig. Cosmids or PCR products from genomic sequence distributed along the length of this contig were used to detect RFLPs between *C. elegans* strains Bristol N2 and Bergerac RC301 (*mgP45* on cosmid C39H11, *mgP46* on cosmid F28D9, *mgP49* on cosmid C35E7, and *mgP50* on cosmid C43H8). These RFLP loci were mapped relative to the *lin-11 daf-16 unc-75* genetic interval. Thirty-three *Daf* non-*Unc* recombinants from a *daf-16(m27) unc-75(e950); daf-2(e1370)/daf-16(+RC301) unc-75(+RC301)* strain and four *Daf* non-*Lin* recombinants from the strain *lin-11(n566) daf-16(m27); daf-2(e1370)/lin-11(+RC301) daf-16(+RC301)* were recovered. In the *Daf* non-*Unc* class, 13/33 recombinant chromosomes carried the Bergerac RC301 of RFLP *mgP46*, 2/30 recombinant chromosomes carried the Bergerac RC301 allele of *mgP49*, and 0/11 and 0/2 carry the Bergerac RC301 allele of RFLP *mgP45* and *mgP50*. In the *Daf* non-*Lin* class, 3/4 of the recombinant chromosomes carried the Bergerac RC301 of RFLP *mgP45* and 0/4 of the recombinant chromosomes carried the Bergerac RC301 allele of *mgP49*. *daf-16(mgDf50); daf-2(e1370)* animals were injected with 100 ng μ l⁻¹ pRF4 and 5 ng μ l⁻¹ of either cosmid R13H8 or B0205. The extent of the deficiencies associated with *daf-16(mgDf47)* and *daf-16(mgDf50)* were further characterized by PCR and Southern blotting. To characterize point mutations, genomic DNA from *daf-16(mg53)*, *daf-16(mg54)*, *daf-16(m26)*, *daf-16(mg87)* and other *daf-16* mutant alleles was PCR-amplified and directly sequenced.

Assaying dauer formation. Characterization of *daf-16* null and *daf-3* null suppression of *daf-2*: synchronized egg broods were grown at 25 °C and scored ~48 h later. Numbers represent totals from 7 or 8 trials for each genotype performed on two different days. Assays for dauer formation on pheromone

were done as described³⁰. Genetic synergy between *daf-1* and *daf-2*: synchronized egg broods were scored at ~48 h and rescored at ~72 h for 25 °C trials, at about 66 and 90 h for both 18 °C and 20 °C trials, or at ~95 and 120 h for 15 °C trials. Numbers in Table 1 represent the summary of five trials of each genotype in two experiments performed on different days.

Expression of *daf-16*. A *daf-16a::GFP* fusion was constructed in the following manner. The *daf-16a* promoter region containing ~6.0 kb of genomic DNA upstream of the *daf-16* 5' UTR and including the 5' UTR and most of exon 1 from *daf-16* was PCR-amplified using primers containing the appropriate restriction sites. This genomic fragment was digested with *Sma*I and *Bam*HI and ligated to pPD95.75 (gift from A. Fire) digested with *Sma*I and *Bam*HI. The *daf-16a::GFP* plasmid was injected at 50 ng μ l⁻¹, together with pRF4 at 100 ng μ l⁻¹.

Received 20 August; accepted 2 October 1997.

- Kimura, K. D. *et al.* *daf-2*, an insulin receptor-like gene that regulates longevity and diapause in *Caenorhabditis elegans*. *Science* **277**, 942–946 (1997).
- Morris, J. Z., Tissenbaum, H. A. & Ruvkun, G. A phosphatidylinositol-3-OH kinase family member regulating longevity and diapause in *Caenorhabditis elegans*. *Nature* **382**, 536–539 (1996).
- Riddle, D. L. Genetic and environmental regulation of dauer larva development in *C. elegans* II, (eds Riddle, D. L. *et al.*) 739–768 (Cold Spring Harbor Press, NY, 1997).
- Kenyon, C. *et al.* A *C. elegans* mutant that lives twice as long as wild type. *Nature* **366**, 461–464 (1993).
- Larsen, P. L., Albert, P. S. & Riddle, D. L. Genes that regulate both development and longevity in *Caenorhabditis elegans*. *Genetics* **139**, 1567–1583 (1995).
- Gottlieb, S. & Ruvkun, G. *daf-2*, *daf-16*, and *daf-23*: Genetically interacting genes controlling dauer formation in *C. elegans*. *Genetics* **137**, 107–120 (1994).
- Vowles, J. J. & Thomas, J. H. Genetic analysis of chemosensory control of dauer formation in *Caenorhabditis elegans*. *Genetics* **130**, 105–123 (1992).
- Patterson, G. I. *et al.* A Smad protein that acts antagonistically in the *C. elegans* TGF- β dauer regulatory pathway. *Genes Dev.* (in the press).
- Clark, K. L. *et al.* Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. *Nature* **364**, 412–420 (1993).
- Galili, N. *et al.* Fusion of a fork head domain to PAX-3 in the solid tumor alveolar rhabdomyosarcoma. *Nature Genet.* **5**, 230–235 (1993).
- Borkhardt, A. *et al.* Cloning and characterization of AFX, the gene that fuses to MLL in acute leukemias with a t(X;11)(q13;q23). *Oncogene* **14**, 195–202 (1997).
- Federicks, W. J. *et al.* The PAX3-FKHR fusion protein created by the t(2;13) translocation in alveolar rhabdomyosarcomas in a more potent transcriptional activator than PAX3. *Mol. Cell. Biol.* **15**, 1522–1535 (1995).
- Kalebic, T., Tsokos, M. & Helman, L. J. *In vivo* treatment with antibody against IGF-1 receptor suppresses growth of human rhabdomyosarcoma and down-regulates p34^{sk2}. *Cancer Res.* **54**, 5531–5534 (1994).
- Shapiro, D. N. *et al.* Antisense-mediated reduction in insulin-like growth factor-I receptor expression suppresses the malignant phenotype of a human alveolar rhabdomyosarcoma. *J. Clin. Invest.* **94**, 1235–1242 (1994).
- Nojima, T. *et al.* A case of alveolar rhabdomyosarcoma with a chromosomal translocation, t(2;13)(q37;q14). *Virchows Arch. Path.* **417**, 357–359 (1990).
- Lai, E. *et al.* HNF-3A, a hepatocyte-enriched transcription factor of novel structure is regulated transcriptionally. *Genes Dev.* **4**, 1427–1436 (1990).
- Kaufmann, E., Muller, D. & Knochel, W. DNA recognition site analysis of *Xenopus* winged helix proteins. *J. Mol. Biol.* **248**, 239–254 (1995).
- Dorman, J. B. *et al.* The *age-1* and *daf-2* genes function in a common pathway to control the lifespan of *Caenorhabditis elegans*. *Genetics* **141**, 1399–1406 (1995).
- Larsen, P. Aging and resistance to oxidative damage in *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* **90**, 8905–8909 (1993).
- Riddle, D. L., Swanson, M. M. & Albert, P. S. Interacting genes in nematode dauer larva formation. *Nature* **290**, 668–671 (1981).
- Toker, A. & Cantley, L. C. Signalling through the lipid products of phosphoinositide-3-OH kinase. *Nature* **387**, 673–676 (1997).
- Tanti, J. F. *et al.* Overexpression of a constitutively active form of phosphatidylinositol 3-kinase is sufficient to promote Glut4 translocation in adipocytes. *J. Biol. Chem.* **271**, 25227–25232 (1996).
- Miller, L. M. *et al.* *lin-31*, a *Caenorhabditis elegans* HNF-3/fork head transcription factor homolog, specifies three alternative cell fates in vulval development. *Genes Dev.* **7**, 933–947 (1993).
- Ren, P. *et al.* Control of *C. elegans* larval development by neuronal expression of a TGF- β homologue. *Science* **274**, 1389–1391 (1996).
- Georgi, L. L., Albert, P. S. & Riddle, D. L. *daf-1*, a *C. elegans* gene controlling dauer larva development, encodes a novel receptor protein kinase. *Cell* **61**, 635–645 (1990).
- Estevez, M. *et al.* The *daf-4* gene encodes a bone morphogenetic protein receptor controlling *C. elegans* dauer larva development. *Nature* **365**, 644–649 (1993).
- Green, J. B., New, H. V. & Smith, J. C. Responses of embryonic *Xenopus* cells to activin and FGF are separated by multiple dose thresholds and correspond to distinct axes of the mesoderm. *Cell* **71**, 731–739 (1992).
- Chen, X., Rubock, M. J. & Whitman, M. A transcriptional partner for MAD proteins in TGF- β signalling. *Nature* **383**, 691–696 (1996).
- O'Brien, R. M. *et al.* Hepatic nuclear factor 3- and hormone-regulated expression of the phosphoenolpyruvate carboxykinase and insulin-like growth factor-binding protein 1 genes. *Mol. Cell. Biol.* **15**, 1747–1758 (1995).
- Golden, J. W. & Riddle, D. L. A pheromone influences larval development in the nematode *Caenorhabditis elegans*. *Science* **218**, 578–580 (1982).

Acknowledgements. We thank S. Chisoe, A. Coulson and the *C. elegans* genome sequencing consortium for sending clones and information and for their help; Y. Liu and F. Lam for technical assistance; Y. Kohara for cDNA clones; R. Barstead for the RB1 and RB2 cDNA libraries; and members of G.R.'s laboratory for discussion and for comments on the manuscript. Some of the strains were provided by the *C. elegans* Genetics Center which is supported by the national Center for Research Resources. This work was supported by a grant from the NIH.

Correspondence and requests for materials should be addressed to G.R. (e-mail: ruvkun@opal.mgh.harvard.edu).

Structure of the cyclin-dependent kinase inhibitor p19^{Ink4d}

Frederich Y. Luh*, Sharon J. Archer†, Peter J. Domaille†, Brian O. Smith*, Darerca Owen*, Deborah H. Brotherton*, Andrew R. C. Raine*, Xu Xu‡, Leonardo Brizuela‡, Stephen L. Brenner† & Ernest D. Laue*

* Cambridge Centre for Molecular Recognition, Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, UK

† Du Pont Merck Pharmaceutical Company, Box 80328, Wilmington, Delaware 19880-0328, USA

‡ Mitotix Inc., One Kendall Square, Building 600, Cambridge, Massachusetts 02139, USA

In cancer, the biochemical pathways that are dominated by the two tumour-suppressor proteins, p53 and Rb, are the most frequently disrupted. Cyclin D-dependent kinases phosphorylate Rb to control its activity and they are, in turn, specifically inhibited by the Ink4 family of cyclin-dependent kinase inhibitors (CDKIs) which cause arrest at the G1 phase of the cell cycle. Mutations in Rb, cyclin D1, its catalytic subunit Cdk4, and the CDKI p16^{Ink4a}, which alter the protein or its level of expression, are all strongly implicated in cancer. This suggests that the Rb 'pathway' is of particular importance¹. Here we report the structure of the p19^{Ink4d} protein, determined by NMR spectroscopy^{2–4}. The structure indicates that most mutations to the p16^{Ink4a} gene, which result in loss of function, are due to incorrectly folded and/or insoluble protein⁵. We propose a model for the interaction of Ink4 proteins with D-type cyclin-Cdk4/6 complexes that might provide a basis for the design of therapeutics against cancer.

The sequences of the Ink4 family of CDKIs are highly conserved (Fig. 1): there are at present four known family members, p15^{Ink4b}, p16^{Ink4a}, p18^{Ink4c} and p19^{Ink4d}. In contrast to the p21/p27/p57 family of CDKIs, which inhibit a broad range of CDKs, the Ink4 family is specific for Cdk4 and Cdk6 (refs 4, 6). The Ink4 proteins, which consist of four or more ankyrin repeats⁷, are expressed in distinct tissue-specific patterns, suggesting that although they have essentially indistinguishable biochemical properties⁴, they are not strictly redundant¹. p15^{Ink4b} is induced in human keratinocytes in response to the transforming growth factor TGF- β (ref. 8). Consistent with its role as a potent inhibitor of Rb phosphorylation, p16^{Ink4a} is altered in familial melanomas⁹, and Ink4a^{-/-} mice spontaneously develop tumours¹⁰. In addition, methylation of the 5' CpG island is associated with transcriptional silencing of p16^{Ink4a} in human cancers¹¹. Both p18^{Ink4c} and p19^{Ink4d} are periodically expressed in proliferating macrophages (maximally during the S-phase of the cell cycle), suggesting that they might limit cyclin-D-dependent kinase activity once cells exit the G1 phase⁴.

Intact mouse p19^{Ink4d} protein was purified from *Escherichia coli* following coexpression with the chaperone proteins GroEL and GroES¹². The recombinant p19^{Ink4d} protein was able to inhibit cyclin D1/Cdk4 at similar concentrations to p16^{Ink4a} (ref. 13, and data not shown). We found that p16^{Ink4a} and p19^{Ink4d} were equally defective in binding an R24C mutant of Cdk4 (ref. 14, and data not shown). These results are consistent with others suggesting that the two proteins have indistinguishable biochemical activities⁴. However, p19^{Ink4d} aggregated less and gave better NMR spectra than p16^{Ink4a} at the high concentrations required for structural studies. ¹H, ¹³C and ¹⁵N chemical shifts of p19^{Ink4d} were assigned after recording three- and four-dimensional (3D, 4D) double and triple resonance NMR spectra of ¹⁵N- and ¹³C/¹⁵N-labelled proteins. The elements of secondary structure (Fig. 1) were identified from the C α and C β

chemical shifts, patterns of nuclear Overhauser effects (NOEs) and $^3J_{\text{HNH}\alpha}$ couplings; they are very similar to the secondary structure determined previously¹⁵ for human p19^{Ink4d}. The 3D structure of the protein was determined using simulated annealing calculations from a total of 3,973 experimental distance and dihedral angle restraints. Identification of unambiguous NOE distance restraints for the preliminary structure calculations was greatly aided by the preparation of perdeuterated and selectively protonated protein samples and will be described in detail elsewhere (F.Y.L. *et al.*, manuscript in preparation). The structures are well defined (Fig. 2a) apart from the first seven residues at the amino terminus and the last three residues at the carboxy terminus, which are flexible in solution, as judged by NMR studies of ^{15}N relaxation in the backbone amides (data not shown).

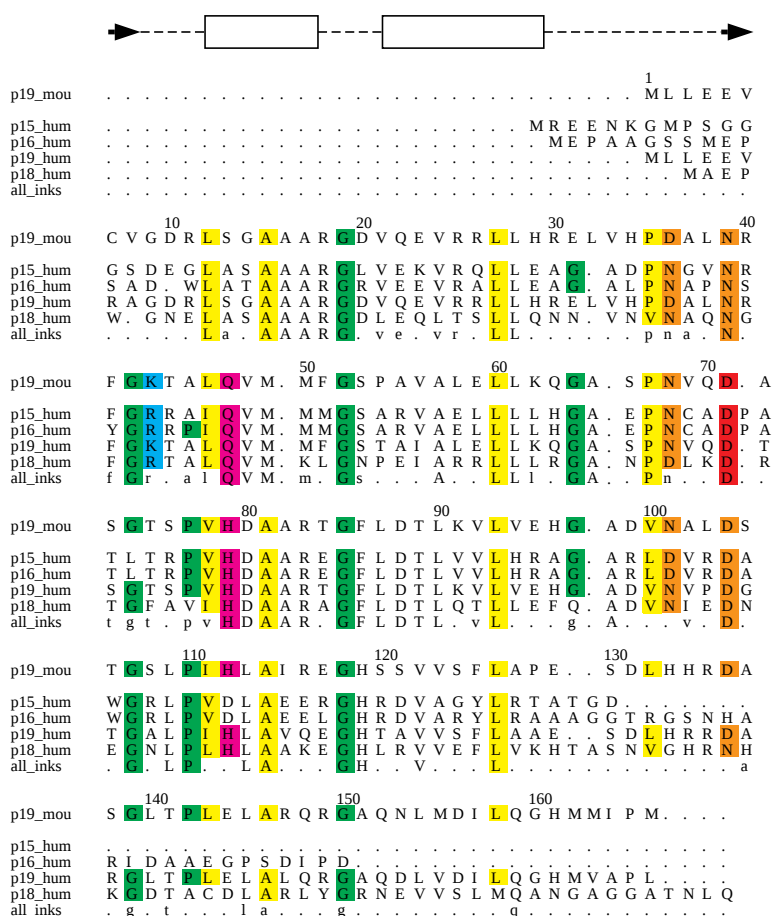
The 3D structure of p19^{Ink4d} is shown in Figs 2 and 3. The protein consists of five ankyrin repeats, whose basic structure and overall packing give an extended L-shaped molecule similar to that of the ankyrin-repeat domain of the p53-binding protein 2 (53BP2)¹⁶ (Fig. 3a). Each ankyrin repeat⁷ comprises an extended strand followed by a helix–loop–helix (HLH) motif and a second extended strand containing a two-residue bulge. Each repeat contributes two residues at both the N and C termini to a series of β -turns which link consecutive repeats. The helices in each HLH pack against each other in an antiparallel fashion to form the long arm of the L, and the extended strand and β -turns form its base.

In general, the β -turns interact with those from adjacent repeats to form a continuous antiparallel β -sheet and adjacent HLHs form four-helix bundles in both proteins. However, in p19^{Ink4d}, the first helix in the second ankyrin is replaced by a mainly extended strand which contains just a single turn of helix (residues 46–50), and the β -turns between ankyrins I–II and II–III are held apart by the side chains on their opposed edges. The changes in packing produced by these differences result in a pronounced kink in the structure,

exemplified by the larger-than-average angle between the first helices of ankyrins I and III. In both p19^{Ink4d} and 53BP2, the first and last ankyrin repeats are incomplete, lacking either the preceding or following extended strands, respectively. The first helix in the final ankyrin repeat of 53BP2, which forms part of the interface with the adjacent SH3 domain, also consists of a mainly extended strand with a single turn of helix at residues 430–434 (ref. 16; Fig. 3a).

A similar role for many of the conserved residues in the ankyrin repeats is seen in 53BP2 and p19^{Ink4d}, although in general the sequence conservation is higher within the Ink4 family of proteins (typically ~45% identical) than it is between 53BP2 and p19^{Ink4d} (~23% identical). As with ankyrins II and III in 53BP2, ankyrins III and IV in p19^{Ink4d} are stabilized by interactions between a conserved histidine in the first helix and the backbone at the base of the following β -turn. In ankyrin II of p19^{Ink4d}, Gln 47 appears to play a similar role (Fig. 2b). In both p19^{Ink4d} and 53BP2, conserved aspartic acid/asparagine residues at the first position of the β -turn stabilize the structure by contacting the backbone of the residue immediately after the turn. In addition in p19^{Ink4d}, Asp 71, which is invariant in the Ink4 family, may interact with Lys 43 (there is always a lysine or arginine at this position) to replace the usual backbone–backbone interactions between β -turns. This is supported by circular dichroism studies of an aspartate-to-asparagine position 71 mutant of p16^{Ink4a} (p19^{Ink4d} numbering) which is unstructured in solution⁵. The side chain of Asp 71 lies on the surface of the protein, so the mutation should not be deleterious in the absence of the proposed interaction with Lys 43 (Fig. 2b). However, similar stabilizing interactions are not present in ankyrin I and could not exist in ankyrin V, which is incomplete. A second group of conserved aspartic acid/asparagine residues at the second position in the bulge may contribute to inter-repeat stabilization by contacting the backbone of the extended strand in the following repeat (Fig. 2b). In both 53BP2 and p19^{Ink4d}, glycines are found at the

Figure 1 Amino-acid sequence of the mouse p19^{Ink4d} protein showing the homology between different ankyrin repeats (defined as in ref. 7, but aligned according to the structure) and between the different members of the Ink4 family of cyclin-dependent kinase inhibitors. Arrows and rectangles indicate the approximate positions of the β -strands and α -helices, respectively. Selected conserved residues are coloured yellow (core hydrophobic), orange (aspartate and asparagine), green (glycine and proline), red (Asp 71), magenta (histidine) and blue (Lys 43) (see text for details).



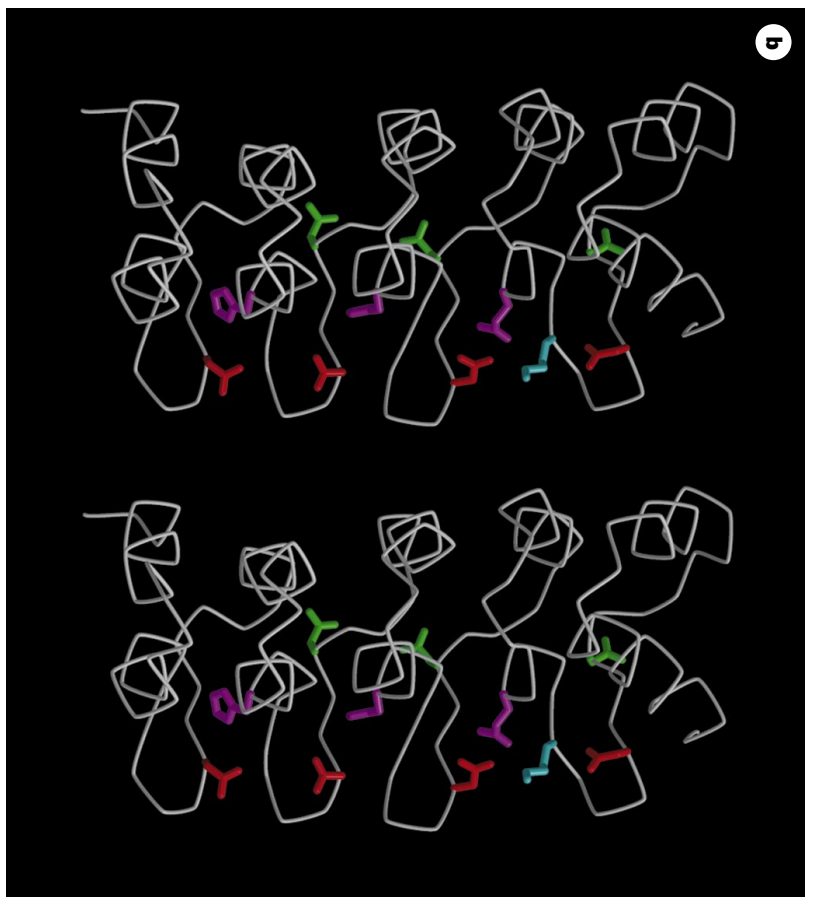
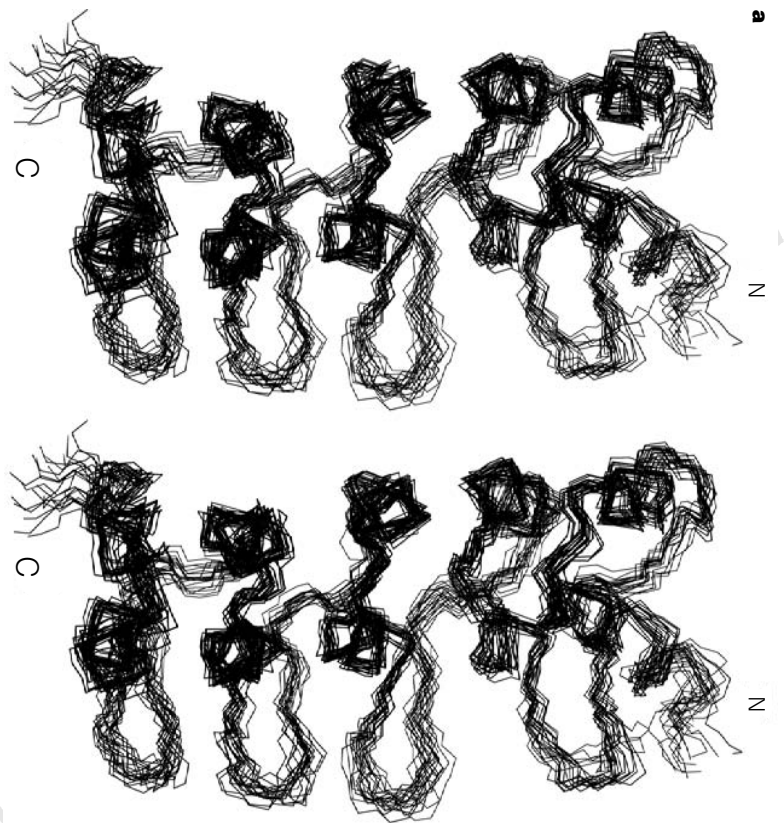


Figure 2 a. Stereoviews of the backbone (N, C α and C') of residues 8–163 from the 20 lowest-energy structures of p19^{h4d} (out of 33 that converged from the 64 computed). In the final structures, the average r.m.s. differences from the mean structure are 0.87 Å (± 0.15) for the backbone atoms, and 1.37 Å (± 0.15) for all the non-hydrogen atoms, respectively. No distance restraint was violated by more than 0.5 Å and no dihedral angle restraint by more than 5.0°. The structures have good covalent geometry and non-bonded contacts (Table 1). **b.** The structure closest to the mean, showing the side chains of key conserved residues that stabilize the structure of the protein. Equivalent positions in the different ankyrin repeats are coloured: histidines, magenta (positions 47 (Gln), 79 and 112); aspartate/asparagine residues in the bulge, green (positions 36, 68 and 101); aspartate/asparagine residues in the β -sheet, red (positions 39, 71, 104 and 136); and Lys 43, cyan. **c.** Same structure as in **b**, but showing key conserved hydrophobic residues. Equivalent positions in the different ankyrin repeats are in blue (positions 12, 46, 78, 111 and 143), red (positions 15, 81, 114 and 146), green (positions 27, 60, 93, 126 and 158) and yellow (positions 35, 67, 100 and 132).

beginning of each ankyrin repeat as the fourth residue of the inter-repeat β -turn, where they allow the chain to twist out of the plane of the turn towards the following helix. Likewise, glycines are found at the beginning of the turn between the two helices in each ankyrin repeat. Glycines are also found at the end of the second helix in ankyrins II and III, where they initiate the turn to the β -sheet (Fig. 1). Prolines, or alanine residues with an equivalent conformation, typically initiate the first helix in the ankyrin repeats in both 53BP2 and p19^{Ink4d} (Fig. 1). There are also many structurally important interactions in both 53BP2 and p19^{Ink4d} between hydrophobic residues in the helices in adjacent ankyrin repeats that pack together in the four-helix bundles (Figs 1 and 2c). The importance of some of these residues for stabilizing the protein structure is supported by circular dichroism and NMR studies of mutants of p16^{Ink4a} which show that, in addition to Asp 71 already mentioned, Pro 77, Gly 97 and Pro 110 (p19^{Ink4d} numbering) are all important for proper folding or maintenance of the protein's structure^{5,17}.

The conservation of amino acids in the ankyrin repeats is also found in the other Ink4 proteins (Fig. 1). In addition, the lengths and positions of the α -helices in p16^{Ink4a} (ref. 17) are very similar to those in p19^{Ink4d} (ref. 15, and our work). In particular, the first helix in ankyrin II appears to be missing in p16^{Ink4a} (ref. 17). Overall, and consistent with the very similar biochemical properties⁴, the results indicate that the different Ink4 family members will have very similar tertiary structures. The structure of p19^{Ink4d} thus allows us to interpret the large amount of mutational data that now exists for p16^{Ink4a}. The gene encoding p16^{Ink4a} is somatically altered in a wide variety of cancers, but there is a more specific pattern of malignancy associated with germline mutations⁹. The strongest link between mutations in the p16^{Ink4a} gene and altered or defective function of the protein can thus be made for inherited germline mutations that are found in familial melanomas. The biochemical features of several of these, as well as of mutants found in primary tumours and cell lines, have been studied^{13,18–21}. Analysis of the structure and the biochemical data suggests that, surprisingly, all the mutations of p16^{Ink4a} that affect binding to Cdk4 and inhibition of cyclin D1/Cdk4 kinase are also likely to affect the folding and/or stability of the protein. Some of these have been discussed, but both our results and

Table 1 Structural statistics for the final ensemble of 20 refined structures of p19^{Ink4d}

R.m.s. deviations	(SA)*	(SA) _c †
From the experimental restraints:		
NOE distances (Å)	0.016 ± 0.0016	0.015
Dihedral angles (°)	0.48 ± 0.10	0.36
From idealized geometry:		
Bonds (Å)	0.0013 ± 0.00006	0.0011
Angles (°)	0.29 ± 0.01	0.28
Final energy		
E_{LJ} ‡ (kJ mol ⁻¹)	−338 ± 29	−384
Assessment of quality according to the Ramachandran plot	All structures	(SA) _c †
Most favoured region	71%	69%
Additionally allowed region	94%	95%

* Represents the average r.m.s. deviations for the ensemble.

† Represents values for the final structure that is closest to the mean.

‡ The Lennard-Jones potential was not used at any stage in the refinement.

those of others⁵ indicate mutations to His 79, Arg 83, Val 94, Ala 98, Asp 104, His 112, Val 122 and Val 123 (p19^{Ink4d} numbering) are all likely to have structural consequences. Our structure suggests that defective protein folding and/or aggregation is the principal means by which p16^{Ink4a} is inactivated in cancer, as demonstrated for some of the mutants^{5,17}.

In Fig. 3a, the p16^{Ink4a} peptide fragments that partially mimic the activity of the intact protein (ref. 22, and S. Wick, M. Dubay and L.B., unpublished results) are shown. These p16^{Ink4a} peptides (residues 80–99; p19^{Ink4d} numbering) are seen to correspond to the HLH motif in ankyrin III. In these peptides, alanine scanning mutagenesis shows that two of the residues that most influence the interaction with Cdk4/6 are at positions 91 and 92 (ref. 22) and these are exposed on the surface of the protein from the face of the second helix (Fig. 3a). These results indicate that this face of the protein might be the site of interaction with the kinase (Fig. 3b). Many of the residues that are conserved in the Ink4 family of CDKIs, but not in ankyrin-repeat proteins generally⁷, are found at positions 80, 83, 87, 88, 89, 92 and 100 within these peptides.

p16^{Ink4a} binds non-competitively to the cyclin D1/Cdk4 kinase with both ATP and the substrate Rb¹³. Consistent with this, we find

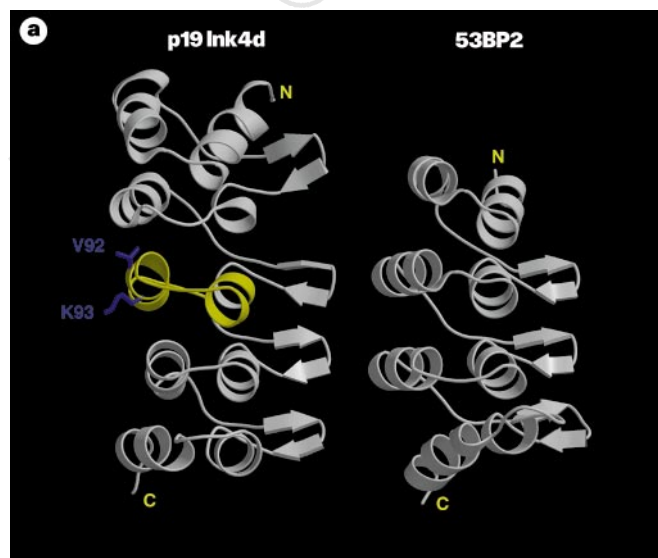
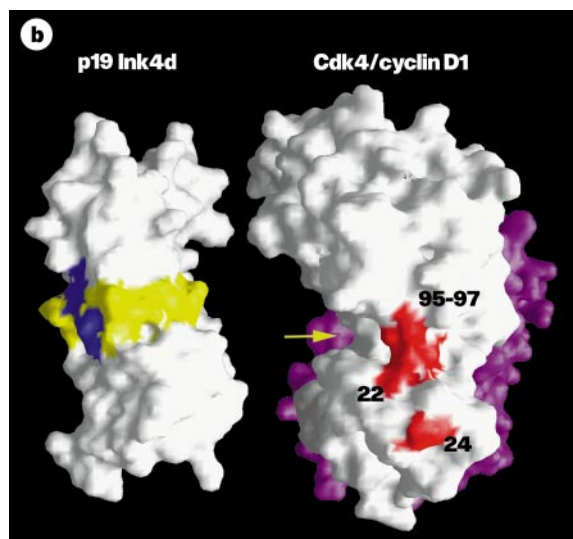


Figure 3 a, Comparison of the structure of p19^{Ink4d} closest to the mean with that of 53BP2. The structures are shown in the same orientation as those in Fig. 2 and the α carbons of structurally equivalent residues in ankyrins III and IV from p19^{Ink4d} and II and III from p53BP2 have been superimposed. The surface-exposed residues of the sequence corresponding to the p16^{Ink4a} peptide fragment (residues 80–99), previously shown partially to mimic the activity of the intact protein, are shown in yellow (ref. 22, and S. Wick, M. Dubay and L.B., unpublished results). The side chains of residues 91 and 92, which, based on alanine scanning



mutagenesis in the peptide²², most influence the interaction with Cdk4 and Cdk6, are shown in dark blue. **b**, Protein surfaces of the same p19^{Ink4d} structure (but now rotated ~55 Å about the y-axis compared to that shown in **a** and Fig. 2) and a model of Cdk4/cyclin D1 (B.O.S., unpublished). Cdk4 is shown in white, cyclin D1 in purple, and Lys 22, Arg 24 and residues 95–97 of Cdk4 are in red; a yellow arrow indicates the position of the active site in Cdk4. These plots were generated using MOLSCRIPT²⁹ and GRASP³⁰.

that Rb, p16^{Ink4a} and cyclin D1/Cdk4 can form a quaternary complex *in vitro* (data not shown). These results suggest that p16^{Ink4a} binds to a site on Cdk4/6 that is different from that for ATP, the substrate Rb and cyclin D1. Similarly, p19^{Ink4d} binds to the cyclin D/Cdk4 complex without disrupting the interaction between the cyclin and Cdk4⁴. In addition, an R24C mutant of Cdk4, which inhibits binding of p16^{Ink4a} and is targeted by cytolytic T lymphocytes in a human melanoma, has been identified⁴. This mutation affects the interaction of p16^{Ink4a}, but not that of p21 or p27, suggesting that the two classes of CDKI bind to different sites. Another mutation that affects p16^{Ink4a} binding is found at Lys 22, and replacement of residues 95–97 in human Cdk4 with those from Cdk2 also abolishes binding to p16^{Ink4a} (M. Serrano and D. Beach, unpublished results). If it is assumed that Cdk4 and Cdk6 have structures similar to Cdk2 (Cdk4 and Cdk6 share 39 and 43% identity with Cdk2, respectively), then all these mutations are likely to cluster together near the junction between the N- and C-terminal lobes of the CDK, next to the active site. The results suggest that the Ink4 class of CDK inhibitors bind in this cleft (Fig. 3b).

The structure of p19^{Ink4d} allows us to interpret and analyse the large amount of data now available on mutants of p16^{Ink4a}. In particular, it indicates which of these mutants might have altered function, as opposed to altered structure, and which would thus qualify for further investigation. Our analysis of p16^{Ink4a} and Cdk4 mutants and of p16^{Ink4a} peptides provides a model for the interaction of the Ink4 class CDKs with Cdk4 and Cdk6 which can now be tested using site-directed mutagenesis and biochemical analysis. □

Methods

Protein purification. Mouse p19^{Ink4d} as a fusion to the glutathione-S-transferase (GST) protein⁴ was coexpressed with GroEL and GroES¹¹ and uniformly ¹³C/¹⁵N- and ¹⁵N-labelled proteins were prepared for the NMR experiments by growing *E. coli* BL21 (DE3) (Novagen), harbouring the expression plasmids, in a minimal medium containing (¹⁵NH₄)₂SO₄ either with or without ¹³C₆-glucose. Deuterated proteins were prepared as described²³. The protein was affinity-purified on glutathione-Sepharose beads (Pharmacia). p19^{Ink4d} was then cleaved from the fusion protein while it was still attached to the resin, using thrombin (Boehringer-Mannheim), which was subsequently removed from the supernatant with an anti-thrombin resin (Sigma). The protein was exchanged into phosphate buffer (20 mM, pH 7.5) containing 100 mM NaCl, 1 mM EDTA and 5 mM DTT.

NMR spectroscopy. All NMR spectra were acquired at 30 °C on Bruker AMX600 NMR spectrometers. The ¹H, ¹³C and ¹⁵N resonances of the backbone of the protein were assigned using a combination of 3-D CBCA(CO)NNH and CBCANHH experiments. The side-chain signals were assigned using 4D HCC(CO)NNH and ¹³C/¹⁵N-separated NOESY experiments, 3D HCCH-COSY and HCCH-TOCSY experiments, and 2D ¹H-¹⁵N HMBC, ¹H-¹³C HSQC, CβHδ and CβHe experiments²⁴.

Structure calculations. The structures were calculated using the program X-PLOR²⁵. Initial structures were based on 1,967 completely unambiguous ¹H-¹H distance restraints obtained from 4D ¹³C/¹⁵N- or ¹⁵N/¹⁵N- and 3D ¹⁵N- or ¹³C-separated NOESY spectra²⁶ of either uniformly ¹⁵N-, ¹³C/¹⁵N- and ²H/¹³C/¹⁵N-labelled or selectively protonated proteins²³. In addition, 82 φ dihedral angle restraints were obtained from a 3D HNHA experiment²⁷. A total of 2,212 ambiguous NOE restraints were then included in the structure calculations. The initial structures were iteratively refined using ARIA methodology²⁸, in which the previous set of structures is used to identify restraint violations and to reduce the level of ambiguity in the ambiguous restraint tables by discarding restraints that contribute little to the intensity.

Received 27 May; accepted 29 July 1997.

- Sherr, C. J. Cancer cell cycles. *Science* **274**, 1672–1677 (1996).
- Chan, F. K. M., Zhang, J., Cheng, L., Shapiro, D. N. & Winoto, D. A. Identification of human and mouse p19, a novel CDK4 and CDK6 inhibitor with homology to p16^{Ink4a}. *Mol. Cell. Biol.* **15**, 2682–2688 (1995).
- Guan, K. *et al.* Isolation and characterization of p19^{Ink4d}, a p16 related inhibitor specific to CDK6 and CDK4. *Mol. Cell. Biol.* **17**, 57–70 (1996).
- Hirai, H., Roussel, M. F., Kato, H.-Y., Ashmun, R. A. & Sherr, C. J. Novel INK4 proteins, p19 and p18, are specific inhibitors of the cyclin D-dependent kinases CDK4 and CDK6. *Mol. Cell. Biol.* **15**, 2672–2681 (1995).

- Zhang, B. & Peng, Z.-y. Defective folding of mutant p16^{Ink4a} proteins encoded by tumor-derived alleles. *J. Biol. Chem.* **271**, 28734–28737 (1996).
- Serrano, M., Hannon, G. J. & Beach, D. A new regulatory motif in cell-cycle control causing specific inhibition of cyclin D/Cdk4. *Nature* **366**, 704–707 (1993).
- Bork, P. Hundreds of ankyrin-like repeats in functionally diverse proteins: mobile modules that cross phyla horizontally. *Prot. Struct. Funct. Genet.* **17**, 363–374 (1993).
- Hannon, G. J. & Beach, D. p15^{Ink4B} is a potential effector of TGF-β-induced cell-cycle arrest. *Nature* **371**, 257–261 (1994).
- Foulkes, W. D., Flanders, T. Y., Pollock, P. M. & Hayward, N. K. The CDKN2A (p16) gene and human cancer. *Mol. Med.* **3**, 5–20 (1997).
- Serrano, M. *et al.* Role of the INK4a locus in tumor suppression and cell mortality. *Cell* **85**, 27–37 (1996).
- Merlo, A. *et al.* 5' CPG island methylation is associated with transcriptional silencing of the tumor-suppressor p16^{CDKN2/MTS} in human cancers. *Nature Med.* **1**, 686–692 (1995).
- Yasukawa, T. *et al.* Increase of solubility of foreign proteins in *Escherichia coli* by co-production of the bacterial thioredoxin. *J. Biol. Chem.* **270**, 25328–25331 (1995).
- Wick, S. T., Dubay, M. M., Imanil, I. & Brizuela, L. Biochemical and mutagenic analysis of the melanoma tumor-suppressor gene-product p16. *Oncogene* **11**, 2013–2019 (1995).
- Wölfel, T. *et al.* A p16^{Ink4a}-insensitive CDK4 mutant targeted by cytolytic T-lymphocytes in a human melanoma. *Science* **269**, 1281–1284 (1995).
- Kalus, W. *et al.* NMR structural characterization of the CDK inhibitor p19^{Ink4d}. *FEBS Lett.* **401**, 127–132 (1997).
- Gorina, S. & Pavletich, N. P. Structure of the p53 tumor suppressor bound to the ankyrin and SH3 domains of 53BP2. *Science* **274**, 1001–1005 (1996).
- Tevelev, A. *et al.* Tumor suppressor p16^{Ink4a}: Structural characterization of wild-type and mutant proteins by NMR and circular dichroism. *Biochemistry* **35**, 9475–9487 (1996).
- Yang, R., Gombart, A. F., Serrano, M. & Koeffler, H. P. Mutational effects on the p16^{Ink4a} tumor suppressor protein. *Cancer Res.* **55**, 2503–2506 (1995).
- Ranade, K. *et al.* Mutations associated with familial melanoma impair p16^{Ink4a} function. *Nature Genet.* **10**, 114–116 (1995).
- Koh, J., Enders, G. H., Dynlacht, B. D. & Harlow, E. Tumour-derived p16 alleles encoding proteins defective in cell-cycle inhibition. *Nature* **375**, 506–510 (1995).
- Parry, D. & Peters, G. Temperature-sensitive mutants of p16^{CDKN2} associated with familial melanoma. *Mol. Cell. Biol.* **16**, 3844–3852 (1996).
- Fähreus, R., Paramio, J. M., Ball, K. L., Lain, S. & Lane, D. P. Inhibition of pRB phosphorylation and cell-cycle progression by a 20-residue peptide derived from p16^{CDKN2/INK4a}. *Curr. Biol.* **6**, 84–91 (1996).
- Smith, B. O. *et al.* An approach to global fold determination using limited NMR data from larger proteins selectively protonated at specific residue types. *J. Biomol. NMR* **8**, 360–368 (1996).
- Clowes, R. T., Crawford, A., Raine, A. R. C., Smith, B. O. & Laue, E. D. Structural studies of proteins using NMR spectroscopy. *Curr. Opin. Biotechnol.* **6**, 81–88 (1995).
- Brünger, A. T. X-PLOR, Version 3.1: A System for X-ray Crystallography and NMR (Yale Univ. Press, New Haven and London, 1992).
- Clow, G. M. & Gronenborn, A. M. Structures of larger proteins in solution—3-dimensional and 4-dimensional heteronuclear NMR spectroscopy. *Science* **252**, 1390–1399 (1991).
- Vuister, G. W. & Bax, A. Quantitative J correlation: A new approach for measuring homonuclear three-bond J(¹H²H³) coupling constants in ¹⁵N-enriched proteins. *J. Am. Chem. Soc.* **115**, 7772–7777 (1993).
- Nilges, M., Marcias, M. J., O'Donoghue, S. I. & Oschkinat, H. Automated NOESY interpretation with ambiguous distance restraints: the refined NMR solution structure of the pleckstrin homology domain from β-spectrin. *J. Mol. Biol.* **269**, 408–422 (1997).
- Kraulis, P. J. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950 (1991).
- Gilson, M. K., Sharp, K. A. & Honig, B. Calculation of the total electrostatic energy of a macromolecular system—solvation energies, binding energies and conformational analysis. *J. Comput. Chem.* **9**, 327–335 (1988).

Acknowledgements. We thank C. J. Sherr for the plasmid expressing GST-p19^{Ink4d}, S. Ishii for plasmids expressing GroEL and GroES; W. Boucher for computer programming; M. Nilges for protocols for the structure calculations; J. Krywko for the model of Cdk4; M. Serrano and D. Beach for assaying p19^{Ink4d} and for communicating unpublished results; N. Pavletich for the coordinates of 53BP2; and A. Murzin for helpful discussion. This work was supported by a grant from the BBSRC. The Cambridge Centre for Molecular Recognition is supported by the BBSRC and the Wellcome Trust.

Correspondence should be addressed to E.D.L. (e-mail: e.d.laue@bioc.cam.ac.uk). The co-ordinates have been deposited in the Brookhaven PDB (accession number 1ap7).

erratum

Chromatin-remodelling factor CHRAC contains the ATPases ISWI and topoisomerase II

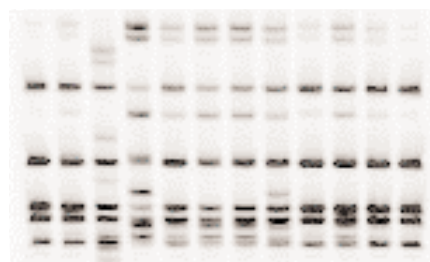
Patrick D. Varga-Weisz, Matthias Wilm, Edgar Bonte, Katia Dumas, Matthias Mann & Peter B. Becker

Nature **388**, 598–602 (1997)

The meaning of the sentence that describes Fig. 5 of this Letter was distorted by erroneous duplication of text and should read: "Comparison of the ATPase activities of CHRAC with that of purified *Drosophila* topoisomerase II (Fig. 5) showed that topoisomerase II was only stimulated by free DNA, whereas CHRAC was stimulated by free DNA as well as by nucleosomes." □

From amplification to antisense

Among this week's batch of products of interest to the molecular biologist are an amplification system for *in-situ* hybridization detection and improved antisense technology.



A sample of the data generated through analysis of human mitochondrial DNA using the Bess T-Scan kit from Epicentre Technologies.

Bess T-Scan kit

From Epicentre Technologies

New technology for the detection of mutations and other changes in the primary structure of DNA fragments

This is a tool for scanning amplified DNA for the presence of base substitutions, deletions, expansions and insertions. It is said to detect greater than 95 per cent of the mutations present within a genetic region. The technology incorporates a simple, single-step enzymatic digestion of PCR-amplified DNA, in which deoxythymidine is randomly replaced with deoxyuridine, creating defined DNA fragments that are resolved on a standard sequencing gel. Shifts in banding patterns of the digested DNA permit easy visualization of primary structure changes. Epicentre says the system does not require extensive optimization, DNA sequencing, special gels, gel electrophoresis conditions, purification, heteroduplex formation or special primers.

Reader Enquiry No. 100

GenPoint

From Dako

*Amplification system for *in-situ* hybridization detection*

Based on the company's catalysed signal amplification (CSA) methodology, GenPoint has been created for use with biotinylated probes. According to Dako, it offers high sensitivity and convenience with none of the complexities and contamination concerns associated with target amplification methods, such as PCR. For example, single copies of HPV16 in SiHa cells can be detected in just five hours, reports Dako. The system can be used with routinely processed tissue (including formalin-fixed paraffin sections) or cell preparations. A similar CSA system for immunohistochemistry offers a signal amplification 20 to 200 times greater than traditional avidin-biotin methods, says the company.

Reader Enquiry No. 101

Anti-DNA-PK

From Kamiya Biomedical

Anti-human DNA-dependent protein kinase (DNA-PK) monoclonal antibodies — three monoclonal antibodies, as well as a cocktail of all three, for research on DNA damage and repair

These clones recognize different epitopes of the catalytic subunit of the 460 K human DNA-dependent protein kinase. DNA-PK phosphorylates a variety of DNA-binding proteins, including transcription factors Spl, Oct1 and p53. DNA-PK is also involved in DNA replication, recombination and double-strand break repair. The Ku autoantigen serves as the DNA-binding component of DNA-PK and is essential for its function. A cocktail of all three clones recognizes DNA-PK from human, mouse, hamster, *Xenopus* and chicken, and is useful for gel-shift assays, immunoprecipitation and western blots.

Reader Enquiry No. 102

Transfection kits

From Sigma

Calcium phosphate and DEAE-dextran kits have been formulated to facilitate mammalian transfection

These new kits for introducing DNA into primary mammalian cells are based on the two most widely used transfection techniques. The kits include all necessary and detailed protocols for transfecting mammalian cells. The calcium phosphate transfection kit may be used for transient or stable transfection in attached or suspension cell cultures. With this technique, DNA is mixed with CaCl_2 and a phosphate buffer to form a fine precipitate that is dispersed over the cultured cells. This precipitate adheres to the cell surface and is taken into the cells. The DEAE-dextran kit is designed primarily for transient transfection of adherent and suspension cell lines. With this kit, positively charged DEAE-dextran complexes tightly with negatively charged DNA. The complex sticks to cell membranes and the entire microparticle is taken in by endocytosis.

Reader Enquiry No. 103

Fusion proteins

From Pierce Chemical

Precoated plates and immobilized ligands for fusion protein screening and purification

Among the most popular fused proteins are the glutathione S-transferase (GST), the maltose binding protein (MBP), histidine-tagged (HisTag) and the biotin fusion pro-

teins. Pierce has developed precoated microtitre plates and other products to screen and purify fusion proteins and their resulting populations. To immobilize fusion protein populations selectively, the company offers microtitre plates precoated with glutathione, dextrin, Ni^{2+} or NeutrAvidin. After adding an aliquot of a fusion protein population to a coated plate and rinsing, the fusion protein is rapidly immobilized, purified and prepared for quantitation using standard ELISA methods. After screening the desired protein, researchers can look to Pierce's collection of immobilized ligands, including glutathione for GST, dextrin for MBP and monomeric avidin for biotin, to facilitate purification. After immobilization, the fusion protein can be eluted in the fused state, or a protease can be introduced to cleave the fusion protein at the appropriate site. For fusion protein complex detection, often achieved using ELISA methods, there is monoclonal anti-GST for the detection of GST fusion proteins and India HisProbe-HRP for the detection of HisTag proteins.

Reader Enquiry No. 104

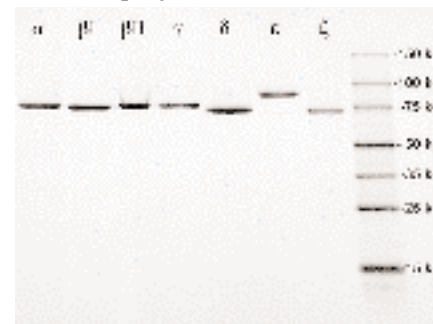
Purified recombinant human PKC isozymes

From PanVera

Recombinant human protein kinase C isozymes α , βI , βII , γ , δ , ϵ and ξ

All isozymes are expressed in a baculovirus system and purified to 95 per cent purity. They are said to offer high specific activity and the appropriate isozyme-specific lipid and calcium dependence. Potential uses include labelling assays, inhibitor screening, elucidation of natural substrates and drug discovery. Alternative uses may include molecular interaction studies and use as gel or western standards. Isozymes are available alone or in a sample pack. The company also offers a non-radioactive kinase assay, Mesacup, and isozyme-specific antibodies.

Reader Enquiry No. 105



PanVera offers a range of recombinant human protein kinase C isozymes.

Fragile X Chemi kits

From Oncor

Fragile X researchers can now move away from radioisotopes with the introduction of two new assays from Oncor

The Fragile X Chemi kit provides digoxigenin-labelled Fragile DNA probes and markers, which can be detected by chemiluminescence following Southern Blot 49 using the SureBlot ChemiHyb and detection kit. Additional benefits include longer probe shelf life, sensitivity comparable with ^{32}P , reduced exposure time and the ability to strip and reprobe blots. Productivity is also said to be increased. The kit contains sufficient reagents for hybridization and chemiluminescent detection of 3,000 cm^2 of membrane, while Fragile X Chemi can process 1,500 cm^2 of membrane.

Reader Enquiry No. 106

Universal GeneComb kit

From Bio-Rad

A kit that is said to give a positive result within an assay time of 30 min, detecting levels of post-amplified DNA down to 10 ng

In a GeneComb assay, the sample is first amplified and deposited in the wells of a microwell strip. Part way along each of the eight nylon teeth that fit into these wells is a 'capture zone' to which a complementary oligonucleotide target probe is fixed. When the comb is dipped into the wells, the sample migrates up to a capture zone, where it is forced into close contact with the probe. If target DNA is present, the resulting hybridization takes place within about 15 min at 37°C. After being dipped into two further microwell strips (one to conjugate with streptavidin/alkaline phosphatase, and another in which the conjugated alkaline phosphatase enzymatically converts a chromogenic reagent), a positive result is indicated with a blue dot on the tooth. Quality is said to be ensured on each comb by a 'control tooth', spotted with both positive (HLA) and

negative control probes. The kit is user-configurable. A capture probe can be bound to the comb by spotting it, and cross-linking to the tooth under ultraviolet light. Two or more probes can be spotted on each tooth to create a multiplex assay, for example. Each kit is sufficient for 32 single assays, with four GeneComb cards and two microplates, each containing two sets of four strip wells. All chromogenic substrates, colour development reagents, running solutions and control oligonucleotides are included.

Reader Enquiry No. 107

Modified antisense oligonucleotides

From Scandinavian Gene Synthesis

Antisense technology continues to develop as a tool in the search for therapies for diseases and is aiding the study of gene function

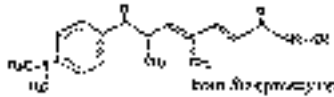
These duplexes interfere with any one of several steps from gene to protein production. Gene expression is altered by the antisense oligonucleotide through the occlusion of processes, requiring a single stranded substrate; stimulation of specific nucleases; direct inactivation by intercalation or cleavage; and prevention of other interactions vital to expression. The design criteria of the antisense oligonucleotide include choice of target sequence and control oligonucleotide, hybridization properties, delivery mechanism and use of nuclease resistant oligonucleotides. Scandinavian Gene Synthesis, GMP-certified manufacturer of high-quality, highly pure DNA, can synthesize a specific sequence with a choice of modifications to suit a given application, either as a straight synthesis or, in combination with other chemistries, as chimeric sequences, in amounts from milligrams to tens of grams.

Reader Enquiry No. 108

These notes are compiled by Brendan Horton from information provided by the manufacturers.

For more details, fill in the reader service card bound inside the journal.

Trichostatin A



Catalog No. 204-11891

This histone deacetylase inhibitor is highly effective at low concentrations, with high specificity. For more information, please call Wake BioProducts: (800) 992-9256; (804) 271-7677.

READER ENQUIRY NO. 98

Bacterial Yeast Growth Media



CONFIDENCE
CONSISTENCY
ACCURACY
RELIABILITY

250 ml Media is available in carboys, powder or pouch form.

Contact BBL 101 or your local distributor for ordering & information
(800) 424-8101 • Fax: (781) 686-0116
Web: www.bbl.com • Email: info@bbl.com

READER ENQUIRY NO. 20

DNA SEQUENCING OLIGONUCLEOTIDES GENOTYPING

By research labs - For research labs

Call us for more information

Molecular Biology Services
King's College School of Medicine and Dentistry
Tel: +44 171 346 3126
Fax: +44 171 733 3877
www.kcl.ac.uk/links/mbs

READER ENQUIRY NO. 19

NEWCASTLE UNIVERSITY MOLECULAR BIOLOGY UNIT



PROTEIN SEQUENCING

Membrane-bound, solid or solution samples

£12 per cycle

5 cycles minimum - no set up charge

For further information contact Dr. Jon Gray
Phone: +44 191 222 8612
Fax: +44 191 222 8129
Email: JOEL.GRAY@newcastle.ac.uk

READER ENQUIRY NO. 14

NEW LOW PRICES - CALL TODAY!

Custom Southern DNA Libraries

- BAC, PAC, P1, COSMID OR OTHER LARGE INSERT CLONES ACCEPTABLE
- RANDOM B13 OR JUST 16 PLANNED LIBRARY PROVIDED
- HIGH QUALITY; >80% INSERT RATIO
- 1-2 WEEK TURNAROUND TIME
- CUSTOMER SATISFACTION GUARANTEED

AND SEQUENCING AND MINIPREP SERVICES AVAILABLE

SECURITY

800-720-4363

5075 Highway, Suite 400
Houston, TX 77056
TEL (713) 829-4363
FAX (713) 829-4322
www.securitylab.com
securitylab@sec.lab.com

READER ENQUIRY NO. 17

VOLUME DNA SEQUENCING

High-throughput single pass sequencing

Full-length sequencing expertise

No royalties or licensing fees

Customer satisfaction guaranteed

NEW LOW PRICES - CALL TODAY!

SECURITY

800-720-4363

5075 Highway, Suite 400
Houston, TX 77056
TEL (713) 829-4363
FAX (713) 829-4322
www.securitylab.com
securitylab@sec.lab.com

READER ENQUIRY NO. 18